

# EXPONENTIAL MARTINGALES AND CHANGES OF MEASURE FOR COUNTING PROCESSES

ALEXANDER SOKOL AND NIELS RICHARD HANSEN

**ABSTRACT.** We give sufficient criteria for the Doléans-Dade exponential of a stochastic integral with respect to a counting process local martingale to be a true martingale. The criteria are sufficiently weak to be useful and verifiable, as illustrated by several non-trivial examples, without introducing artificial constraints. In particular, they make it possible to construct nonexplosive point processes with intensities adapted to a general filtration by a change of measure.

## 1. Introduction

The motivation for this paper is the problem of constructing nonexplosive dynamic processes via a change of measure on the background probability space. The objective is to derive verifiable conditions in a counting process context for the exponential martingale to be a true martingale without introducing artificial constraints. As discussed recently by (Gjessing et al., 2010), it is of general interest to formulate a statistical model of a dynamic counting process in terms of a family of candidate intensities, and it is then essential to be able to verify that the intensities give well defined nonexplosive models. To this end we need conditions on the candidate intensities. If the intensity is adapted to the filtration generated by the counting process itself precise results are obtainable by transferring the problem to a canonical setup, see Theorem 5.2.1 in (Jacobsen, 2005). If  $N$  is a counting process, which, under  $P$ , is a homogeneous Poisson counting process, a combination of the mentioned theorem and Exercise 4.4.5 in (Jacobsen, 2005) gives the following result; for an intensity process  $\lambda$  such that

$$(1.1) \quad \lambda_t \leq a(N_{t-})$$

for a sequence  $a(n)$  satisfying  $\sum_{n=1}^{\infty} \frac{1}{a(n)} = \infty$  there is a measure  $Q$  with Radon-Nikodym derivative with respect to  $P$  being an exponential martingale such that  $N$  is a nonexplosive counting process with intensity  $\lambda$  under  $Q$ . The result is mentioned in (Gjessing et al., 2010) as the Jacobsen condition. It holds a priori on the canonical spaces considered in (Jacobsen, 2005). It is generally not possible to lift a measure from a canonical space to an abstract space but if the intensity is adapted to the filtration

---

2000 *Mathematics Subject Classification.* Primary 60G44; Secondary 60G55.

*Key words and phrases.* Counting Process, Martingale, Exponential martingale, Girsanov, Intensity, Uniform integrability, Absolute continuity.

generated by  $N$  the fact that the exponential martingale is a true martingale and not just a local martingale can be lifted. This shows that (1.1), which is a very weak condition, is sufficient for the exponential martingale to be a true martingale if the intensity is adapted to the filtration generated by  $N$ . This, in turn, allows for the construction of a counting process on bounded intervals with intensity  $\lambda$  by a change of measure if (1.1) holds.

Alternative approaches to ensure the existence of a nonexplosive counting process with a given intensity are surveyed in (Gjessing et al., 2010), but a weak and general but yet verifiable condition that the exponential local martingale is a true martingale is missing. Results by Lépingle and Mémin, (Lépingle and Mémin, 1978), are mentioned but not applied, a restrictive Novikov-type condition is mentioned, and the most general, explicit condition mentioned in (Gjessing et al., 2010) is (25). This is a growth condition on  $\lambda_t^\alpha$  with  $\alpha > 1$ , which is typically too strong or difficult to verify in practice. A consequence of our results is that a growth condition with  $\alpha = 1$  is sufficient, which is much more useful.

Our starting point is the paper by Lépingle and Mémin, (Lépingle and Mémin, 1978), and their general results, which we adapt to the specific case of Doléans-Dade exponentials of stochastic integrals with respect to counting process local martingales. This is, in itself, not enough to obtain sufficiently weak criteria – in several examples such a specialization would still leave artificial constraints on the intensities considered. We circumvent this by a trick that essentially allows for a restriction to arbitrarily small time intervals, and doing so we can remove the otherwise artificial constraints.

We illustrate how the criteria can be verified. We consider, in particular, examples of interacting diffusion and jump processes for which the general framework is suitable.

## 2. Summary of results

Consider a filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$  satisfying the usual conditions, see (Protter, 2005), Section I.1 for the definition of this as well as other standard probabilistic concepts. We say that  $N$  is a nonexplosive  $d$ -dimensional counting process if  $N$  is càdlàg and piecewise constant with jumps of size one, and no coordinates of  $N$  jump at the same time. We say that a process  $X$  is locally bounded if there is a sequence of stopping times increasing almost surely to infinity such that  $X^{T_n} 1_{(T_n > 0)}$  is bounded. Let  $\lambda$  be a nonnegative, predictable and locally bounded  $d$ -dimensional process. Then  $\lambda$  is almost surely integrable on compacts with respect to the Lebesgue measure. We say that  $N$  is a counting process with intensity  $\lambda$  if it holds that  $N_t^i - \int_0^t \lambda_s^i ds$  is a local martingale for each  $i$ . Note in particular that since the predictable  $\sigma$ -algebra considered is the one generated by the filtration  $(\mathcal{F}_t)_{t \geq 0}$ , the intensity is allowed to depend on other processes than just  $N$ .

We recall the definition of Doléans-Dade exponentials. Assume given a semimartingale  $X$  with initial value zero. By (Protter, 2005), Theorem II.37, the stochastic differential equation  $Z_t = 1 + \int_0^t Z_{s-} dX_s$  has a càdlàg adapted solution, unique up to indistinguishability, and the solution is

$$(2.1) \quad \mathcal{E}(X)_t = \exp\left(X_t - \frac{1}{2}[X^c]_t\right) \prod_{0 < s \leq t} (1 + \Delta X_s) \exp(-\Delta X_s),$$

where  $X^c$  is the continuous martingale part of  $X$ , see p. 209 of (He et al., 1992) for the definition of the continuous martingale part of a semimartingale. We call  $\mathcal{E}(X)$  the Doléans-Dade exponential of  $X$ . Assume that  $\Delta X \geq -1$ , it then holds that  $\mathcal{E}(X)$  is always nonnegative. Furthermore, defining  $R = \inf\{t \geq 0 \mid \Delta X_t = -1\}$ , Theorem I.4.61 of Jacod and Shiryaev (2003) then shows that  $\mathcal{E}(X)_t$  is positive for  $0 \leq t < R$  and zero for  $R \leq t < \infty$ . Therefore, we may always write

$$(2.2) \quad \mathcal{E}(X)_t = 1_{(t < R)} \exp\left(X_t - \frac{1}{2}[X^c]_t + \sum_{0 < s \leq t} \log(1 + \Delta X_s) - \Delta X_s\right).$$

If  $X$  is a local martingale,  $\mathcal{E}(X)$  is a local martingale as well, and in this case, we refer to  $\mathcal{E}(X)$  as an exponential martingale.

Now assume given a  $d$ -dimensional counting process  $N$  with nonnegative, predictable and locally bounded intensity  $\lambda$ , and assume given another  $d$ -dimensional nonnegative, predictable and locally bounded process  $\mu$ .

**Definition 2.1.** We say that  $\mu$  is  $\lambda$ -compatible if it holds that  $\mu_t^i(\omega) = 0$  whenever  $\lambda_t^i(\omega) = 0$ , and if the process  $\gamma$  defined by  $\gamma_t^i = \mu_t^i(\lambda_t^i)^{-1}$  for  $i \leq d$  is locally bounded.

In Definition 2.1, we use the convention that zero divided by zero is equal to one. Now assume that  $\mu$  is  $\lambda$ -compatible. Define  $M$  to be the  $d$ -dimensional local martingale given by  $M_t^i = N_t^i - \int_0^t \lambda_s^i ds$ . Put  $\gamma_t^i = \mu_t^i(\lambda_t^i)^{-1}$  and  $H_t^i = \gamma_t^i - 1$  for  $t \geq 0$ . As we have assumed that  $\mu$  is  $\lambda$ -compatible,  $\gamma$  and  $H$  are both well-defined and locally bounded real-valued processes. We define  $(H \cdot M)_t = \sum_{i=1}^d \int_0^t H_s^i dN_s^i$ ,  $H \cdot M$  is then a one-dimensional process. The local boundedness of  $H$  ensures that  $H \cdot M$  is well-defined. Defining  $\log_+ x = \max\{0, \log x\}$  for  $x \geq 0$ , with the convention that the logarithm of zero is minus infinity, our main results are the following.

**Theorem 2.2.** Assume that  $\lambda$  and  $\mu$  are nonnegative, predictable and locally bounded. Assume that  $\mu$  is  $\lambda$ -compatible. It holds that  $\mathcal{E}(H \cdot M)$  is a martingale if there is an  $\varepsilon > 0$  such that whenever  $0 \leq u \leq t$  with  $|t - u| \leq \varepsilon$ , one of the following two conditions are satisfied:

$$(2.3) \quad E \exp\left(\sum_{i=1}^d \int_u^t (\gamma_s^i \log \gamma_s^i - (\gamma_s^i - 1)) \lambda_s^i ds\right) < \infty \quad \text{or}$$

$$(2.4) \quad E \exp\left(\sum_{i=1}^d \int_u^t \lambda_s^i ds + \int_u^t \log_+ \gamma_s^i dN_s^i\right) < \infty.$$

**Corollary 2.3.** *Assume that  $\lambda = 1$  and assume that  $\mu$  is nonnegative, predictable and locally bounded. Then  $\mu$  is  $\lambda$ -compatible. It holds that  $\mathcal{E}(H \cdot M)$  is a martingale if there is an  $\varepsilon > 0$  such that whenever  $0 \leq u \leq t$  with  $|t - u| \leq \varepsilon$ , one of the following two conditions are satisfied:*

$$(2.5) \quad E \exp \left( \sum_{i=1}^d \int_u^t \mu_s^i \log_+ \mu_s^i ds \right) < \infty \quad \text{or} \quad E \exp \left( \sum_{i=1}^d \int_u^t \log_+ \mu_s^i dN_s^i \right) < \infty.$$

The immediate use of Theorem 2.2 and its corollary is as an existence result for nonexplosive counting processes with particular intensities, as the change of measure obtained from the martingale property of  $\mathcal{E}(H \cdot M)$  yields the existence of a nonexplosive counting process distribution with given intensity  $\mu$  on a bounded time interval  $[0, t]$ . That this is the case may be seen from Lemma 3.5, discussed below, which shows that under the measure  $Q_t$  with Radon-Nikodym derivative  $\mathcal{E}(H \cdot M)_t$  with respect to  $P$ ,  $N$  is a counting process with intensity  $1_{[0, t]} \mu + 1_{(t, \infty)} \lambda$ .

As a specific application, let us assume that we are interested in constructing a statistical model for a nonexplosive counting processes. We assume given a filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$  and a  $d$ -dimensional counting process  $N$  such that under  $P$ ,  $N^i$  has intensity  $\lambda_t^i = 1$ . Fix a timepoint  $t$  and let us assume that we are interested in considering a statistical model on the time interval  $[0, t]$  based on a family of intensities  $(\mu_\theta)_{\theta \in \Theta}$ . If  $\mu_\theta$  satisfies the criteria of Corollary 2.3, we find that  $\mathcal{E}(H \cdot M)$  is a martingale, and so  $\mathcal{E}(H \cdot M)_t$  has unit mean. Letting  $Q_\theta$  be the probability measure with Radon-Nikodym derivative  $\mathcal{E}(H \cdot M)_t$  with respect to  $P$ , it holds that under  $Q_\theta$ ,  $N$  is a counting process with intensity, and the intensity is  $\mu_\theta$  on  $[0, t]$ . Furthermore, the family  $(Q_\theta)_{\theta \in \Theta}$  is dominated by  $P$ , and the likelihood function is known in explicit form. Thus, Corollary 2.3 has allowed us to construct the statistical model and prove that explosion does not occur.

As regards checking the criteria in practice, an important property to note is that the criteria only need to be checked locally, in the sense that it is only necessary to find some  $\varepsilon > 0$  such that the criteria holds for  $0 \leq u \leq t$  with  $|t - u| \leq \varepsilon$ . This seemingly innocent property makes it possible to apply the criteria in several interesting situations. In particular, it allows us to extend the criterion (25) of (Gjessing et al., 2010) from  $\alpha > 1$  to  $\alpha \geq 1$ .

The rest of the paper is organized as follows. In Section 3, we present the prerequisites for the main results. In Section 4, we give some examples of applications of the results. Appendix A contains proofs.

### 3. Prerequisites for the main results

In this section, we present the prerequisites for our main results, Theorem 2.2 and Corollary 2.3. We first recall some known results on exponential martingales. Lemma 3.1 yields some basic information on exponential martingales, and Lemma 3.2 is a simple criterion for the martingale property of exponential martingales. The results are folklore, see for example p. 140 of (Protter, 2005) for the continuous case, and we therefore do not give proofs.

**Lemma 3.1.** *If  $M$  is a local martingale with  $\Delta M \geq -1$  and initial value zero,  $\mathcal{E}(M)$  is a nonnegative local martingale and a supermartingale,  $E\mathcal{E}(M)_t \leq 1$  and  $\mathcal{E}(M)_\infty$  always exists as an almost sure limit with  $E\mathcal{E}(M)_\infty \leq 1$ .*

**Lemma 3.2.** *Let  $M$  be a local martingale with  $\Delta M \geq -1$  and initial value zero.  $\mathcal{E}(M)$  is a uniformly integrable martingale if and only if  $E\mathcal{E}(M)_\infty = 1$ , and  $\mathcal{E}(M)$  is a martingale if and only if  $E\mathcal{E}(M)_t = 1$  for all  $t \geq 0$ .*

For  $M$  a local martingale with  $\Delta M \geq -1$  and initial value zero, the question of when  $\mathcal{E}(M)$  is a uniformly integrable martingale or a true martingale has been treated many times in the literature, see for example (Novikov, 1972), (Kazamaki and Sekiguchi, 1983), (Kazamaki, 1994) and (Cherny and Shiryaev, 2001) for results in the case of continuous  $M$ , and (Lépingle and Mémin, 1978), (Izumisawa et al., 1979) and (Kallsen and Shiryaev, 2002) for results in the general case. In this paper, we will apply the criteria obtained in (Lépingle and Mémin, 1978) to integrals of compensated counting processes. The two main results from that article are the following, where  $\Pi_p^*$  denotes the dual predictable projection, see Definition 5.21 of (He et al., 1992).

**Theorem 3.3** ((Lépingle and Mémin, 1978), Theorem III.1). *Let  $M$  be a local martingale with initial value zero and  $\Delta M \geq -1$ . Let  $R = \inf\{t \geq 0 \mid \Delta M_t = -1\}$ . Define  $B$  by putting  $B_t = \frac{1}{2}[M^c]_{t \wedge R} + \sum_{0 < s \leq t \wedge R} (1 + \Delta M_s) \log(1 + \Delta M_s) - \Delta M_s$ . If  $B$  is locally integrable and  $\exp(\Pi_p^* B_\infty)$  is integrable, then  $\mathcal{E}(M)$  is a uniformly integrable martingale.*

**Theorem 3.4** ((Lépingle and Mémin, 1978), Theorem III.7). *Let  $M$  be a local martingale with initial value zero and  $\Delta M > -1$ . Define a finite variation process  $A$  by putting  $A_t = \frac{1}{2}[M^c]_t + \sum_{0 < s \leq t} \log(1 + \Delta M_s) - \frac{\Delta M_s}{1 + \Delta M_s}$ . If  $\exp(A_\infty)$  is integrable, then  $\mathcal{E}(M)$  is a uniformly integrable martingale.*

Note that the function  $x \mapsto (1+x) \log(1+x) - x$  is well-defined on  $(-1, \infty)$  and has limit 1 for  $x$  tending to  $-1$  from above. Therefore, the function can be continuously extended to  $[-1, \infty)$ , and so the process  $B$  mentioned in Theorem 3.3 is well-defined up to and including time  $R$ .

Now consider given a  $d$ -dimensional nonexplosive counting process  $N$  with nonnegative, predictable and locally bounded intensity  $\lambda$  as well as another nonnegative, predictable and locally bounded process  $\mu$  which is  $\lambda$ -compatible. As in the previous section,  $M$  is the  $d$ -dimensional local martingale defined by  $M_t^i = N_t^i - \int_0^t \lambda_s^i ds$ . Furthermore, we

also use the notation that  $\gamma^i = \mu_t^i(\lambda_t^i)^{-1}$  and  $H_t^i = \gamma_t^i - 1$ . Recall that the assumption that  $\mu$  is  $\lambda$ -compatible implies that both  $\gamma$  and  $H$  are locally bounded. Integrals are vector integrals in the sense that  $H \cdot M$  denotes the one-dimensional process defined by  $H \cdot M = \sum_{i=1}^d H^i \cdot M^i$ .

We begin by showing that if  $\mathcal{E}(H \cdot M)^T$  is a martingale, changing the measure using  $\mathcal{E}(H \cdot M)_T$  as a Radon-Nikodym derivative corresponds to changing the intensity of  $N$  on  $[0, T]$  from  $\lambda$  to  $\mu$ .

**Lemma 3.5.** *Let  $T$  be a stopping time and assume that  $\mathcal{E}(H \cdot M)^T$  is a uniformly integrable martingale. With  $Q$  being the probability measure with Radon-Nikodym derivative  $\mathcal{E}(H \cdot M)_T$  with respect to  $P$ , it holds that  $N$  is a counting process under  $Q$  with intensity  $1_{[0, T]} \mu + 1_{(T, \infty)} \lambda$ . In particular, if  $\mathcal{E}(H \cdot M)$  is a martingale, it holds for any  $t \geq 0$  and with  $Q_t$  being the probability measure with Radon-Nikodym derivative  $\mathcal{E}(H \cdot M)_t$  with respect to  $P$  that  $N$  is a counting process under  $Q_t$  with intensity  $1_{[0, t]} \mu + 1_{(t, \infty)} \lambda$ .*

Lemma 3.5 shows that given  $\lambda$  and  $\mu$ ,  $\mathcal{E}(H \cdot M)$  is the relevant exponential martingale to consider for changing the distribution of  $N$  from a counting process with intensity  $\lambda$  to a counting process with intensity  $\mu$ , where  $H_t^i = \mu_t^i(\lambda_t^i)^{-1} - 1$ . In general, we cannot expect  $\mathcal{E}(H \cdot M)$  to be a uniformly integrable martingale, only an ordinary martingale, because the distributions of counting processes with intensities which differ sufficiently will in general be singular. For example, the distributions of two homogeneous Poisson processes with different intensities are singular, see Proposition 3.24 of (Karr, 1986).

As an aside, note that the measure  $Q$  obtained in Lemma 3.5 of course always will be absolutely continuous with respect to  $P$ . A natural question to ask is when  $Q$  and  $P$  will be equivalent. This is the case when the Radon-Nikodym derivative is almost surely positive. The following lemma gives a condition for this to be the case.

**Lemma 3.6.** *If the set of zeroes of  $\mu$  has Lebesgue measure zero,  $\mathcal{E}(H \cdot M)$  is almost surely positive.*

The following two lemmas will be used in the proof of our main results. The first lemma allows us to restrict our attention to small deterministic time intervals when proving the martingale property of exponential martingales, and the second lemma decomposes an exponential martingale into the product of two exponential martingales, corresponding to successive changes of intensity from  $\lambda$  to  $\mu$  and  $\mu$  to  $\mu + \nu$ . This will, colloquially speaking, allow us to consider the large and small parts of  $\mu$  separately when proving the martingale property.

**Lemma 3.7.** *Let  $M$  be a local martingale with  $\Delta M \geq -1$ , and let  $\varepsilon > 0$ . If  $\mathcal{E}(M^t - M^u)$  is a martingale whenever  $0 \leq u \leq t$  with  $|t - u| \leq \varepsilon$ , where  $M^t$  denotes the process  $M$  stopped at time  $t$ , then  $\mathcal{E}(M)$  is a martingale.*

**Lemma 3.8.** *Let  $\nu$  be nonnegative, predictable and locally bounded. Assume that  $\mu$  is  $\lambda$ -compatible and that  $\mu + \nu$  is  $\mu$ -compatible. Then  $\mu + \nu$  is also  $\lambda$ -compatible. Put  $(H_\lambda^{\mu+\nu})_t^i = (\mu_t^i + \nu_t^i)(\lambda_t^i)^{-1} - 1$ ,  $(H_\lambda^\mu)_t^i = \mu_t^i(\lambda_t^i)^{-1} - 1$  and  $(H_\mu^{\mu+\nu})_t^i = (\mu_t^i + \nu_t^i)(\mu_t^i)^{-1} - 1$ .*

Define  $d$ -dimensional processes  $M^\lambda$  and  $M^\mu$  by putting  $(M^\lambda)_t^i = N_t^i - \int_0^t \lambda_s^i ds$  and  $(M^\mu)_t^i = N_t^i - \int_0^t \mu_s^i ds$ . It then holds that  $\mathcal{E}(H_\lambda^{\mu+\nu} \cdot M^\lambda) = \mathcal{E}(H_\lambda^\mu \cdot M^\lambda) \mathcal{E}(H_\mu^{\mu+\nu} \cdot M^\mu)$ .

Combining the above lemmas allows us to prove Theorem 2.2. The proof, along with the proofs of the lemmas above, may be found in Appendix A.

#### 4. Examples

In this section, we give examples where the conditions in Theorem 2.2 and Corollary 2.3 may be verified. Our first example shows how Theorem 2.2 under certain circumstances allows for changes of the intensity where the new intensity is an affine function of the old intensity. Such criteria were also discussed in Theorem 2 of (Røysland, 2011), where the new intensity  $\mu$  was assumed to be related to the initial intensity  $\lambda$  by the relationship  $|\mu_t - \lambda_t| \leq \theta \sqrt{\lambda_t}$ .

**Example 4.1.** Assume that  $d$  is equal to one. Assume that there is  $\delta > 0$  such that  $\lambda_s \geq \delta$  and assume that  $\mu_t \leq \alpha + \beta \lambda_s$ . If there is  $\varepsilon > 0$  such that for  $0 \leq u \leq t$  with  $|t - u| \leq \varepsilon$ ,  $\int_u^t \lambda_s ds$  has an exponential moment of order  $(1 + (\alpha\delta^{-1} + \beta) \log_+(\alpha\delta^{-1} + \beta))$ , then  $\mathcal{E}(H \cdot M)$  is a martingale.

In the remainder of the examples, we will assume that  $\lambda = 1$ , such that  $N$  is a  $d$ -dimensional standard Poisson process, and give particular cases where Corollary 2.3 may be applied.

**Example 4.2.** Assume that  $\mu$  is a nonnegative, predictable and locally integrable process, and assume that there is  $\varepsilon > 0$  such that  $\exp(\varepsilon \langle H \cdot M \rangle_t)$  is integrable for all  $t \geq 0$ . In this case, the first criterion of Corollary 2.3 may be applied to show that  $\mathcal{E}(H \cdot M)$  is a martingale.

Example 4.2 is noteworthy because of the following. In (Protter and Shimbo, 2008), applying the results of Lépingle and Mémin (1978), the following Novikov-type criterion is demonstrated: If  $M$  is a locally square integrable local martingale with  $\Delta M \geq -1$  and  $\exp(\frac{1}{2} \langle M^c \rangle_\infty + \langle M^d \rangle_\infty)$  is integrable, then  $\mathcal{E}(M)$  is a uniformly integrable martingale. Here,  $M^c$  and  $M^d$  denote the continuous and purely discontinuous parts of the local martingale, respectively, see Theorem 7.25 of He et al. (1992). Furthermore, Protter and Shimbo (2008) argue by example that the constant 1 in front of  $\langle M^d \rangle_\infty$  cannot in general be exchanged with  $1 - \varepsilon$  for any  $\varepsilon > 0$ . Example 4.2, however, shows that when proving the martingale property instead of the uniformly integrable martingale property, for the particular type of local martingale considered here, the constant 1 may in fact be exchanged with any positive number. This is a consequence of the particular form of  $\langle H \cdot M \rangle$  combined with the fact that we are endeavouring to prove the martingale property and not the uniformly integrable martingale property.

**Example 4.3.** Assume that  $\mu$  is a nonnegative, predictable and locally integrable process satisfying  $\mu_t^i \leq \alpha + \beta \sum_{j=1}^d N_{t-}^j$ . Then both criteria of Corollary 2.3 may be applied to obtain that  $\mathcal{E}(H \cdot M)$  is a martingale.

The existence of counting processes with intensities affinely bounded by the total number of jumps as in Example 4.3 is well known, see Example 4.4.5 of (Jacobsen, 2005). The above example yields the same existence through a measure change on general probability spaces, independent of canonical spaces. This is the extension of criterion (25) of (Gjessing et al., 2010) from  $\alpha > 1$  to  $\alpha \geq 1$  mentioned earlier.

If the process  $\mu$  is exactly affine in the sense that  $\mu_t^i = \alpha + \beta \sum_{j=1}^d N_{t-}^j$ , the martingale property of  $\mathcal{E}(H \cdot M)$  may be obtained by direct calculation. However, this does not in itself imply that the same result holds when we only have  $\mu_t^i \leq \alpha + \beta \sum_{j=1}^d N_{t-}^j$ . In general, such “monotonicity” properties of the martingale property for exponential martingales do not hold, see for example (Kazamaki, 1994), Example 1.13.

Next, we consider two examples involving intensities given as solutions to stochastic differential equations. In both cases, we assume given a Brownian motion relative to the given filtration  $(\mathcal{F}_t)$ , meaning in the  $d$ -dimensional case that  $(W^i)_t^2 - t$  is an  $(\mathcal{F}_t)$  martingale for  $i \leq d$  and  $W_t^i W_t^j$  is an  $(\mathcal{F}_t)$  martingale for  $i, j \leq d$  with  $i \neq j$ . We denote such a process an  $(\mathcal{F}_t)$ -Brownian motion. By Lévy’s characterisation of Brownian motion for general filtered probability spaces, see Theorem IV.33.1 of (Rogers and Williams, 2000b), this requirement ensures that the characteristic properties of the Brownian motion interact well with the filtration  $(\mathcal{F}_t)$ . By  $\mathbb{M}(d, d)$ , we denote the set of  $d \times d$  matrices with real entries.

**Example 4.4.** Consider three mappings  $A : \mathbb{N}_0^d \times \mathbb{R}_+^d \rightarrow \mathbb{R}^d$ ,  $B : \mathbb{N}_0^d \times \mathbb{R}_+^d \rightarrow \mathbb{M}(d, d)$  and  $\sigma : \mathbb{N}_0^d \times \mathbb{R}_+^d \rightarrow \mathbb{M}(d, d)$  such that for all  $\eta \in \mathbb{N}_0^d$ ,  $A(\eta, \cdot)$ ,  $B(\eta, \cdot)$  and  $\sigma(\eta, \cdot)$  are continuous and bounded and such that  $\sigma$  always is positive definite. With  $T_n^i$  denoting the  $n$ ’th jump time for  $N^i$  and  $Z_t^i = t - T_{N_t^i}^i$ , let  $X$  be a solution to the  $d$ -dimensional stochastic differential equation

$$(4.1) \quad dX_t = (A(N_t, Z_t) + B(N_t, Z_t)X_t) dt + \sigma(N_t, Z_t) dW_t$$

with initial value  $x_0$  in  $\mathbb{R}^d$ , where  $W$  is an  $(\mathcal{F}_t)$  Brownian motion independent of  $N$ . Let  $\phi : \mathbb{R}^d \rightarrow \mathbb{R}_+^d$  be Lipschitz and put  $\mu_t = \phi(X_t)$ . Assume that there are  $\delta > 0$  and  $c_A, c_B, c_\sigma > 0$  such that

$$(4.2) \quad \sup_{t \geq 0} \|A(\eta, t)\|_2 \leq c_A \|\eta\|_1^{1-\delta}$$

$$(4.3) \quad \sup_{t \geq 0} \|\sigma(\eta, t)\|_2 \leq c_\sigma \|\eta\|_1^{(1-\delta)/2}$$

$$(4.4) \quad \sup_{t \geq 0} \|B(\eta, t)\|_2 \leq c_B,$$

where  $\|\cdot\|_2$  in the first case denotes the Euclidean norm and in the two latter cases denote the operator norm induced by the Euclidean norm, and  $\|\cdot\|_1$  denotes the  $\mathcal{L}^1$  norm in



$\mathbb{N}_0^d$ . Then, the first criteria of Corollary 2.3 may be applied to obtain that  $\mathcal{E}(H \cdot M)$  is a martingale.

**Example 4.5.** Let  $(\xi_n)_{n \geq 0}$ ,  $(a_n)_{n \geq 0}$  and  $(b_n)_{n \geq 0}$  be sequences in  $\mathbb{R}$ . Assume that  $b_n \neq 0$  for  $n \geq 0$  and assume that  $X$  satisfies the one-dimensional stochastic differential equation

$$(4.5) \quad dX_t = a_{N_t} + b_{N_t} X_t dt + \sigma dW_t + (\xi_{N_t} - X_{t-}) dN_t,$$

with initial value  $\xi_0$  and  $\sigma > 0$ , where  $W$  is an  $(\mathcal{F}_t)$  Brownian motion independent of  $N$ . Put  $\mu_t = |X_{t-}|$ . Assume that there are  $\alpha, \beta > 0$  such that

$$(4.6) \quad |\xi_n| \leq \alpha + \beta n$$

$$(4.7) \quad |a_n/b_n| \leq \alpha + \beta n$$

$$(4.8) \quad |b_n| \leq \alpha.$$

Then, the second criteria of Corollary 2.3 may be applied to obtain that  $\mathcal{E}(H \cdot M)$  is a martingale.

Examples 4.4 and 4.5 show how Corollary 2.3 may be used to construct counting processes with intensities not adapted to the filtration induced by  $N$  itself. Note that by Corollary 11.5.3 of (Shreve, 2004),  $W$  is always independent of  $N$ , so the independence requirements in the above are mentioned only for clarity. Also note that in Example 4.4, the required bounds on the coefficients hold independently of the norms on  $\mathbb{N}_0^d$ ,  $\mathbb{R}^d$  and  $\mathbb{M}(d, d)$  chosen, since all norms on finite-dimensional vector spaces are equivalent. The existence of solutions to the stochastic differential equations are proved in the appendix.

The interpretation of the two examples are the following. In Example 4.4, the intensity is a transformed diffusion process with mean reversion level, mean reversion speed and diffusion coefficient which are deterministic between jumps. A simple example may be obtained as follows. Let  $X$  be a solution to the one-dimensional stochastic differential equation

$$(4.9) \quad dX_t = \beta(\alpha \exp(-\gamma(t - T_{N_t})) - X_t) dt + \sigma dW_t,$$

where  $\alpha, \beta, \gamma \geq 0$  and  $T_n$  is the  $n$ 'th event time of  $N$ . Define  $\mu_t = |X_t|$ .  $\mu$  is then a process of the type given in Example 4.4. Except when  $X$  is nonpositive,  $\mu$  behaves as a diffusion immediately after each jump of  $N$ , with a mean reversion level  $\alpha$ , reverting to this level at rate  $\beta$ , and furthermore, the mean reversion level decreases exponentially with rate  $\gamma$  in  $t - T_{N_t}$ , which is the time since the last jump of  $N$ .

In Example 4.5, the intensity is the absolute value of a linear diffusion process with constant coefficients between jumps. Furthermore, the intensity is reset to the level  $\xi_n$  at the  $n$ 'th jump of  $N$ .

**Example 4.6.** Consider mappings  $\phi_i : \mathbb{R} \rightarrow [0, \infty)$  and  $h_{ij} : [0, \infty) \rightarrow \mathbb{R}$ . Define

$$(4.10) \quad \mu_t^i = \phi_i \left( \sum_{j=1}^d \int_0^{t-} h_{ij}(t-s) dN_s^j \right).$$

If  $\phi^i$  is Borel measurable with  $\phi_i(x) \leq |x|$  and  $h_{ij}$  is bounded, then  $\mathcal{E}(H \cdot M)$  is a martingale.

Example 4.6 yields a change of measure to a probability measure where the counting process is a multidimensional Hawkes process. In general, many specifications of  $\phi$  and  $h$  will yield exploding counting processes and there will exist no measure change yielding the required intensity change.

The above examples all give various types of sufficient criteria for the martingale property of  $\mathcal{E}(H \cdot M)$  using Corollary 2.3. As an aside, we may ask whether the classical necessary and sufficient criterion for nonexplosion for piecewise constant intensities, see Theorem 2.3.2 of (Norris, 1997), may be replicated as a criterion for the martingale property of  $\mathcal{E}(H \cdot M)$ . The following example shows that this is the case.

**Example 4.7.** Let  $d = 1$ , let  $(\alpha_n)$  be a sequence of positive numbers and let  $\lambda_t = \alpha_{N_t-}$ . Then  $\mathcal{E}(H \cdot M)$  is a martingale if and only if  $\sum_{n=0}^{\infty} \frac{1}{\alpha_n}$  is divergent.

## Appendix A. Proofs

### A.1. Proofs for Section 3.

**Lemma A.1.** *Let  $N$  have intensity  $\lambda$ . If  $X$  is a process which is nonnegative, predictable and locally bounded, and it holds almost surely that pathwisely, the set of zeroes of  $X$  has Lebesgue measure zero, then it almost surely holds that the zeroes of  $X$  are disjoint from the jump times of  $N^i$  for all  $i$ .*

**Proof of Lemma A.1.** As  $X$  is predictable, the set of zeroes of  $X$  is a predictable set. Thus, the integral process  $\int_0^t 1_{(X_s=0)} dM_s^i$  is a local martingale. Let  $(T_n)$  be a localising sequence such that  $\int_0^t 1_{(X_s=0)} 1_{(t \leq T_n)} dN_s^i$  is bounded and such that  $\int_0^t 1_{(X_s=0)} 1_{(t \leq T_n)} dM_s^i$  is a true martingale. In particular,  $E \int_0^t 1_{(X_s=0)} 1_{(t \leq T_n)} dM_s^i = 0$ . By our assumptions, it holds almost surely that pathwisely,  $1_{(X_s=0)}$  is zero except on a Lebesgue null set. Therefore,

$$(A.1) \quad E \int_0^t 1_{(X_s=0)} 1_{(t \leq T_n)} dN_s^i = E \int_0^t 1_{(X_s=0)} 1_{(t \leq T_n)} \lambda_s^i ds = 0,$$

where the integrals are well-defined as  $\int_0^t 1_{(X_s=0)} 1_{(t \leq T_n)} dN_s^i$  is bounded. Therefore, as  $\int_0^t 1_{(X_s=0)} 1_{(t \leq T_n)} dN_s^i$  is nonnegative, we conclude that  $\int_0^t 1_{(X_s=0)} 1_{(t \leq T_n)} dN_s^i$  is almost surely zero. Letting  $n$  and then  $t$  tend to infinity, we find that  $\int_0^\infty 1_{(X_s=0)} dN_s^i$  is almost surely zero, and this implies that almost surely, the set of zeroes of  $X$  is disjoint from the jump times of  $N^i$ . As the coordinate  $i$  was arbitrary, the result follows.  $\square$

**Proof of Lemma 3.6.** Note that  $\Delta(H \cdot M)_t = \sum_{i=1}^d H_t^i \Delta N_t^i$ . By Lemma A.1, the set of zeroes of  $\mu^i$  is disjoint from the jump times of  $N^i$ . Therefore, the set of zeroes of  $\gamma^i$

is disjoint from the jump times of  $N^i$  as well, and so the set where  $H^i$  is  $-1$  is disjoint from the jump times of  $N^i$ . We conclude that almost surely,  $H \cdot M$  has no jumps of size  $-1$ . Theorem I.4.61 of Jacod and Shiryaev (2003) then shows that  $\mathcal{E}(H \cdot M)$  is almost surely positive.  $\square$

**Lemma A.2.** *Let  $M$  be a local martingale with  $\Delta M \geq -1$  and let  $T$  be a stopping time. Assume that  $\mathcal{E}(M)^T$  is a uniformly integrable martingale. Let  $Q$  be the probability measure having Radon-Nikodym derivative  $\mathcal{E}(M)_T$  with respect to  $P$ . If  $L$  is a local martingale under  $P$  such that  $[L, M^T]$  is locally integrable under  $P$ , then  $L - \langle L, M^T \rangle$  is a local martingale under  $Q$ , where the angle bracket is calculated under  $P$ .*

*Proof.* First note that as  $Q$  has a density with respect to  $P$ ,  $Q$  is absolutely continuous with respect to  $P$ . With  $Z$  being the likelihood process for  $Q$  with respect to  $P$ , meaning that  $Z_t = E(\frac{dQ}{dP} | \mathcal{F}_t)$ , we have  $Z_t = E(\mathcal{E}(M)_\infty^T | \mathcal{F}_t) = \mathcal{E}(M)_t^T$  up to indistinguishability. In particular,  $Z_0 = 1$  almost surely. By an examination of the proof of the predictable Girsanov theorem, Theorem III.41 of (Protter, 2005), we therefore find that the theorem can be applied in spite of our not having assumed that  $\mathcal{F}_0$  is a sub- $\sigma$ -algebra of the  $P$ -completion of  $\{\emptyset, \Omega\}$ , as the theorem in (Protter, 2005) otherwise requires.

Now consider a process  $L$  which is a local martingale under  $P$  such that  $[L, M^T]$  is locally integrable under  $P$ . Note that  $[L, \mathcal{E}(M^T)] = [L, \mathcal{E}(M^T)_- \cdot M^T] = \mathcal{E}(M^T)_-^T \cdot [L, M^T]$ . As  $\mathcal{E}(M)_-$  is left-continuous, it is locally bounded. Therefore, as  $[L, M^T]$  is locally integrable, the process  $\mathcal{E}(M^T)_-^T \cdot [L, M^T]$  is locally integrable as well. Thus,  $[L, \mathcal{E}(M^T)]$  is locally integrable under  $P$ , so the predictable covariation of this process is well-defined under  $P$ . Then, Theorem III.41 of (Protter, 2005), applies and yields that the process given by  $L_u - \int_0^u \mathcal{E}(M^T)_{s-}^{-1} d\langle \mathcal{E}(M^T), L \rangle_s$  is a  $Q$  local martingale, where the angle bracket is calculated under  $P$ . Noting that

$$\begin{aligned} L_u - \int_0^u \frac{1}{\mathcal{E}(M^T)_{s-}} d\langle \mathcal{E}(M^T), L \rangle_s &= L_u - \int_0^u \frac{1}{\mathcal{E}(M^T)_{s-}} d\langle \mathcal{E}(M^T)_- \cdot M^T, L \rangle_s \\ (A.2) \qquad \qquad \qquad &= L_u - \langle L, M^T \rangle_u, \end{aligned}$$

the result follows.  $\square$

**Proof of Lemma 3.5.** Fix a stopping time  $T$ . By definition,  $Q$  has Radon-Nikodym derivative  $\mathcal{E}(H \cdot M)_T$  with respect to  $P$ . We wish to apply Lemma A.2 in order to prove the result. We first check that  $[M^i, (H \cdot M)^T]$  is locally integrable under  $P$ . Since  $M^i$  has paths of finite variation,  $[M^i, M^j]_t = \sum_{0 < s \leq t} \Delta M_s^i \Delta M_s^j = \sum_{0 < s \leq t} \Delta N_s^i \Delta N_s^j = [N^i, N^j]$ , and in particular  $[M^i] = N^i$ . As the coordinates of  $N$  have no common jumps, we have

$$\begin{aligned} [M^i, (H \cdot M)^T] &= \sum_{j=1}^d H^j 1_{[0, T]} \cdot [M^j, M^i] \\ (A.3) \qquad \qquad \qquad &= \sum_{j=1}^d H^j 1_{[0, T]} \cdot [N^j, N^i] = H^i 1_{[0, T]} \cdot [N^i]. \end{aligned}$$

Because we have assumed that  $H$  is locally bounded, this is locally integrable. From Lemma A.2, we then conclude that  $M^i - \langle M^i, (H \cdot M)^T \rangle$  is a local martingale under  $Q_t$ . Next, we know that under  $P$ ,  $(\Pi_p^* N^i)_t = \int_0^t \lambda_s^i ds$ , and we also know that  $H$  and  $1_{[0,T]}$  are predictable. Therefore,  $\langle M^i, (H \cdot M)^T \rangle_s = \Pi_p^*(H^i 1_{[0,T]} \cdot [N^i])_s = \int_0^s H_u^i 1_{(u \leq T)} \lambda_u^i du$ , which allows us to conclude that

$$\begin{aligned}
 M_s^i - \langle M^i, (H \cdot M)^T \rangle_s &= N_s^i - \int_0^s \lambda_s^i ds - \int_0^s H_u^i 1_{(u \leq T)} \lambda_u^i du \\
 &= N_t^i - \int_0^s \lambda_u^i + \left( \frac{\mu_u^i}{\lambda_u^i} - 1 \right) 1_{(u \leq T)} \lambda_u^i du \\
 (A.4) \qquad &= N_t^i - \int_0^s \mu_u^i 1_{[0,T]}(u) + \lambda_u^i 1_{(T,\infty)}(u) du.
 \end{aligned}$$

This proves that under  $Q$ ,  $N$  has intensity  $1_{[0,T]} \mu + 1_{(T,\infty)} \lambda$ . The results for the case where  $\mathcal{E}(H \cdot M)$  is a martingale then follows by considering stopping times which are constant.  $\square$

**Lemma A.3.** *Let  $M^1, \dots, M^n$  be local martingales with pairwise zero quadratic covariation. Then  $\mathcal{E}(\sum_{k=1}^n M^k) = \prod_{k=1}^n \mathcal{E}(M^k)$ .*

*Proof.* This follows by Theorem II.38 of (Protter, 2005) and an induction proof.  $\square$

**Proof of Lemma 3.7.** Let  $\varepsilon > 0$  be given such that  $\mathcal{E}(M^t - M^u)$  is a martingale whenever  $0 \leq u \leq t$  with  $|t - u| \leq \varepsilon$ . By Lemma 3.2, to show that  $\mathcal{E}(M)$  is a martingale, it suffices to show  $E\mathcal{E}(M)_t = 1$  for all  $t \geq 0$ . As  $\mathcal{E}(M)$  is a supermartingale,  $E\mathcal{E}(M)_t$  is decreasing, and we know that  $\mathcal{E}(M)_0 = 1$ . Therefore, it will suffice to prove  $E\mathcal{E}(M)_{n\varepsilon} = 1$  for  $n \geq 1$ . Now, for naturals  $n < m$  it holds that  $n \leq m - 1$  and so

$$\begin{aligned}
 &[M^{n\varepsilon} - M^{(n-1)\varepsilon}, M^{m\varepsilon} - M^{(m-1)\varepsilon}] \\
 &= [M^{n\varepsilon}, M^{m\varepsilon}] - [M^{n\varepsilon}, M^{(m-1)\varepsilon}] - [M^{(n-1)\varepsilon}, M^{m\varepsilon}] + [M^{(n-1)\varepsilon}, M^{(m-1)\varepsilon}] \\
 (A.5) \qquad &= [M]^{n\varepsilon} - [M]^{n\varepsilon} - [M]^{(n-1)\varepsilon} + [M]^{(n-1)\varepsilon},
 \end{aligned}$$

which is zero. Then,  $\mathcal{E}(M)_{n\varepsilon} = \prod_{k=1}^n \mathcal{E}(M^{k\varepsilon} - M^{(k-1)\varepsilon})_{n\varepsilon} = \prod_{k=1}^n \mathcal{E}(M^{k\varepsilon} - M^{(k-1)\varepsilon})_{k\varepsilon}$  by Lemma A.3. Also, by our assumptions,  $\mathcal{E}(M^{k\varepsilon} - M^{(k-1)\varepsilon})$  is a martingale. Now note

that for  $1 < k \leq n$ ,

$$\begin{aligned}
E \prod_{i=1}^k \mathcal{E}(M^{i\varepsilon} - M^{(i-1)\varepsilon})_{i\varepsilon} &= E \left( E \left( \prod_{i=1}^k \mathcal{E}(M^{i\varepsilon} - M^{(i-1)\varepsilon})_{i\varepsilon} \middle| \mathcal{F}_{(k-1)\varepsilon} \right) \right) \\
&= E \left( E(\mathcal{E}(M^{k\varepsilon} - M^{(k-1)\varepsilon})_{k\varepsilon} | \mathcal{F}_{(k-1)\varepsilon}) \prod_{i=1}^{k-1} \mathcal{E}(M^{i\varepsilon} - M^{(i-1)\varepsilon})_{i\varepsilon} \right) \\
&= E \left( \mathcal{E}(M^{k\varepsilon} - M^{(k-1)\varepsilon})_{(k-1)\varepsilon} \prod_{i=1}^{k-1} \mathcal{E}(M^{i\varepsilon} - M^{(i-1)\varepsilon})_{i\varepsilon} \right) \\
(A.6) \quad &= E \prod_{i=1}^{k-1} \mathcal{E}(M^{i\varepsilon} - M^{(i-1)\varepsilon})_{i\varepsilon},
\end{aligned}$$

since  $\mathcal{E}(M^{k\varepsilon} - M^{(k-1)\varepsilon})_{(k-1)\varepsilon} = 1$ . By induction,  $E\mathcal{E}(M)_{n\varepsilon} = 1$ , and so we conclude that  $\mathcal{E}(M)$  is a martingale.  $\square$

**Proof of Lemma 3.8.** That  $\mu + \nu$  is  $\lambda$ -compatible follows as  $\mu + \nu$  is  $\mu$ -compatible and  $\mu$  is  $\lambda$ -compatible. Furthermore,  $M^\lambda$  and  $M^\mu$  are processes of finite variation, so we find

$$\begin{aligned}
[H_\lambda^\mu \cdot M^\lambda, H_\mu^{\mu+\nu} \cdot M^\mu]_t &= \sum_{0 \leq s \leq t} \Delta(H_\lambda^\mu \cdot M^\lambda)_s \Delta(H_\mu^{\mu+\nu} \cdot M^\mu)_s \\
&= \sum_{i=1}^d \sum_{0 \leq s \leq t} (H_\lambda^\mu)_s^i (H_\mu^{\mu+\nu})_s^i \Delta N_s^i \\
(A.7) \quad &= \sum_{i=1}^d \int_0^t (H_\lambda^\mu)_s^i (H_\mu^{\mu+\nu})_s^i dN_s^i.
\end{aligned}$$

Therefore,  $\mathcal{E}(H_\lambda^\mu \cdot M^\lambda) \mathcal{E}(H_\mu^{\mu+\nu} \cdot M^\mu) = \mathcal{E}(H_\lambda^\mu \cdot M^\lambda + H_\mu^{\mu+\nu} \cdot M^\mu + H_\lambda^\mu H_\mu^{\mu+\nu} \cdot N)$  by Theorem II.38 of (Protter, 2005). We find

$$\begin{aligned}
(A.8) \quad &(H_\lambda^\mu \cdot M^\lambda + H_\mu^{\mu+\nu} \cdot M^\mu + H_\lambda^\mu H_\mu^{\mu+\nu} \cdot N)_t \\
&= \sum_{i=1}^d \int_0^t (H_\lambda^\mu)_s^i + (H_\mu^{\mu+\nu})_s^i + (H_\lambda^\mu)_s^i (H_\mu^{\mu+\nu})_s^i dN_s^i - \int_0^t (H_\lambda^\mu)_s^i \lambda_s^i + (H_\mu^{\mu+\nu})_s^i \mu_s^i ds.
\end{aligned}$$

Now noting that

$$\begin{aligned}
(A.9) \quad &(H_\lambda^\mu)_t^i + (H_\mu^{\mu+\nu})_t^i + (H_\lambda^\mu)_t^i (H_\mu^{\mu+\nu})_t^i \\
&= \left( \frac{\mu_t^i}{\lambda_t^i} - 1 \right) + \left( \frac{\mu_t^i + \nu_t^i}{\mu_t^i} - 1 \right) + \left( \frac{\mu_t^i}{\lambda_t^i} - 1 \right) \left( \frac{\mu_t^i + \nu_t^i}{\mu_t^i} - 1 \right) \\
&= \frac{\mu_t^i \mu_t^i + \nu_t^i}{\lambda_t^i \mu_t^i} - 1 = \frac{\mu_t^i + \nu_t^i}{\lambda_t^i} - 1 = (H_\lambda^{\mu+\nu})_t^i
\end{aligned}$$

as well as  $(H_\lambda^\mu)_t^i \lambda_t^i + (H_\mu^{\mu+\nu})_t^i \mu_t^i = \mu_t^i - \lambda_t^i + \mu_t^i + \nu_t^i - \mu_t^i = \mu_t^i + \nu_t^i - \lambda_t^i$ , we may conclude

$$\begin{aligned} & H_\lambda^\mu \cdot M^\lambda + H_\mu^{\mu+\nu} \cdot M^\mu + H_\lambda^\mu H_\mu^{\mu+\nu} \cdot N \\ (A.10) \quad &= \sum_{i=1}^d \int_0^t (H_\lambda^{\mu+\nu})_s^i dN_s^i - \sum_{i=1}^d \int_0^t (H_\lambda^{\mu+\nu})_s^i \lambda_s^i ds = H_\lambda^{\mu+\nu} \cdot M^\lambda, \end{aligned}$$

yielding the desired result.  $\square$

**Proof of Theorem 2.2.** By Lemma 3.7, it suffices to show that  $\mathcal{E}((H \cdot M)^t - (H \cdot M)^u)$  is a martingale when  $0 \leq u \leq t$  with  $|t - u| \leq \varepsilon$ . Let such a pair of  $u$  and  $t$  be given and let  $L = (H \cdot M)^t - (H \cdot M)^u$ . With  $R$  and  $B$  as in Theorem 3.3, we have for  $r \geq 0$  that

$$\begin{aligned} B_r &= \frac{1}{2}[L^c]_{r \wedge R} + \sum_{0 < s \leq r \wedge R} (1 + \Delta L_s) \log(1 + \Delta L_s) - \Delta L_s \\ (A.11) \quad &= \sum_{i=1}^d \int_0^r 1_{[0,R]}(s) 1_{[u,t]}(s) ((1 + H_s^i) \log(1 + H_s^i) - H_s^i) dN_s^i. \end{aligned}$$

From this, we obtain that  $B$  is locally integrable, and as  $1_{[0,R]}$  is a predictable process, we have

$$\begin{aligned} (\Pi_p^* B)_\infty &= \sum_{i=1}^d \int_0^\infty 1_{[0,R]}(s) 1_{[u,t]}(s) ((1 + H_s^i) \log(1 + H_s^i) - H_s^i) \lambda_s^i ds \\ &= \sum_{i=1}^d \int_u^t 1_{[0,R]}(s) (\gamma_s^i \log \gamma_s^i - (\gamma_s^i - 1)) \lambda_s^i ds \\ (A.12) \quad &\leq \sum_{i=1}^d \int_u^t (\gamma_s^i \log \gamma_s^i - (\gamma_s^i - 1)) \lambda_s^i ds. \end{aligned}$$

Therefore, if the first integrability criterion is satisfied,  $\mathcal{E}(L)$  is a uniformly integrable martingale by Theorem 3.3, in particular a martingale. This proves the first claim.

Next, we consider the case where the second integrability criterion is satisfied. We will use Lemma 3.8 to prove that  $\mathcal{E}((H \cdot M)^t - (H \cdot M)^u)$  is a martingale in this case. To this end, we define predictable  $d$ -dimensional processes  $\mu^-$  and  $\mu^+$  by

$$(A.13) \quad (\mu^-)_t^i = \mu_t^i 1_{(\mu_t^i \leq \lambda_t^i)} + \lambda_t^i 1_{(\mu_t^i > \lambda_t^i)}$$

$$(A.14) \quad (\mu^+)_t^i = (\mu_t^i - \lambda_t^i) 1_{(\mu_t^i > \lambda_t^i)}.$$

We then have  $\mu = \mu^+ + \mu^-$ . Also define  $\gamma^* = \frac{\mu^-}{\lambda}$  and  $\gamma^{**} = \frac{\mu}{\mu^+}$ , and  $H^* = \gamma^* - 1$  and  $H^{**} = \gamma^{**} - 1$ . Now, as  $\lambda$  and  $\mu$  are predictable,  $\mu^-$  and  $\mu^+$  are predictable as well. Furthermore,  $\mu^-$  and  $\mu^+$  are both nonnegative and locally bounded. We claim that  $\mu^-$  is  $\lambda$ -compatible and that  $\mu$  is  $\mu^-$ -compatible.

To show this, first note that as  $\mu$  is  $\lambda$ -compatible, it holds that if  $\lambda^i$  is zero,  $\mu^i$  is zero and so  $(\mu^-)^i$  is zero. Also,  $(\gamma^*)_t^i = \gamma_t^i 1_{(\mu_t^i \leq \lambda_t^i)} + 1_{(\mu_t^i > \lambda_t^i)}$ , so  $\gamma^*$  is locally bounded since  $\gamma$  is locally bounded. Thus,  $\mu^-$  is  $\lambda$ -compatible. Next, note that if  $(\mu^-)^i$  is zero, it either

holds that  $\mu^i$  is zero and  $\mu^i \leq \lambda^i$  or that  $\lambda^i$  is zero and  $\mu^i > \lambda^i$ . The latter is not possible since we have assumed that  $\mu$  is  $\lambda$ -compatible. Therefore, if  $(\mu^-)^i$  is zero,  $\mu_t^i$  is zero. Furthermore,  $(\gamma^{**})_t^i = 1_{(\mu_t^i \leq \lambda_t^i)} + \gamma_t^i 1_{(\mu_t^i > \lambda_t^i)}$ , so  $\gamma^{**}$  is locally bounded as well. Thus,  $\mu$  is  $\mu^-$ -compatible.

Now put  $L^* = (H^* \cdot M)^t - (H^* \cdot M)^u$  and  $L^{**} = (H^{**} \cdot M)^t - (H^{**} \cdot M)^u$ . It holds that  $\mathcal{E}(L) = \mathcal{E}(L^*)\mathcal{E}(L^{**})$ . As  $(\mu^-)^i \leq \lambda^i$ , we conclude that  $-1 \leq (H^*)^i \leq 0$ . We will apply Theorem 3.3 to the local martingale  $L^*$ . By the same calculations as above, noting that  $(1+x)\log(1+x) \leq 0$  when  $-1 \leq x \leq 0$ , we obtain

$$\begin{aligned} (\Pi_p^* B)_\infty &= \sum_{i=1}^d \int_u^t ((1 + (H^*)_s^i) \log(1 + (H^*)_s^i) - (H^*)_s^i \lambda_s^i) ds \\ (A.15) \quad &\leq - \sum_{i=1}^d \int_u^t (H^*)_s^i \lambda_s^i ds \leq \sum_{i=1}^d \int_u^t \lambda_s^i ds. \end{aligned}$$

Since we have assumed that the second integrability condition holds, we obtain

$$(A.16) \quad E \exp \left( \sum_{i=1}^d \int_u^t \lambda_s^i ds \right) \leq E \exp \left( \sum_{i=1}^d \int_u^t \lambda_s ds + \int_u^t \log_+ \gamma_s^i dN_s^i \right) < \infty,$$

so Theorem 3.3 shows that  $\mathcal{E}(L^*)$  is a uniformly integrable martingale. Let  $Q$  be the measure with Radon-Nikodym derivative  $\mathcal{E}(L^*)_\infty$  with respect to  $P$ . By Lemma 3.8,  $E^P \mathcal{E}(L)_\infty = E^Q \mathcal{E}(L^{**})_\infty$ . To show that  $\mathcal{E}(L)$  is a martingale, it suffices to show that this is equal to one. To do so, we will apply Theorem 3.4 to show that  $\mathcal{E}(L^{**})$  is a uniformly integrable martingale under  $Q$ . To this end, we first show that  $L^{**}$  is a local martingale under  $Q$ . Note that  $L^{**}$  and  $L^*$  both have paths of finite variation, and

$$\begin{aligned} [L^{**}, L^*]_\infty &= \sum_{i=1}^d \sum_{u < s \leq t} (H_s^{**})^i (H_s^*)^i \Delta N_s^i \\ (A.17) \quad &= \sum_{i=1}^d \sum_{u < s \leq t} \left( \frac{\mu_s^i}{(\mu^-)_s^i} - 1 \right) \left( \frac{(\mu^-)_s^i}{\lambda_s^i} - 1 \right) \Delta N_s^i = 0, \end{aligned}$$

since for fixed  $s$  and  $i$ , the first factor is zero when  $\mu_s^i < \lambda_s^i$ , and the second factor is zero when  $\mu_s^i \geq \lambda_s^i$ . In particular,  $[L^{**}, L^*]$  is locally integrable under  $P$ , so  $\langle L^{**}, L^* \rangle$  exists under  $P$  and is zero. By Lemma A.2, we find that  $L^{**} - \langle L^{**}, (L^*)^t \rangle$  is a local martingale under  $Q$ . By what we just showed,  $\langle L^{**}, (L^*)^t \rangle$  is zero, so  $L^{**}$  is a local martingale under  $Q$ . Next, note that  $(H^{**})_t^i = (\gamma^{**})_t^i - 1 = 1_{(\mu_t^i \leq \lambda_t^i)} + \gamma_t^i 1_{(\mu_t^i > \lambda_t^i)} - 1 \geq 0$ , so  $\Delta L^{**} \geq 0 > -1$ , and therefore Theorem 3.4 is applicable. Now, with  $A$  as in Theorem

3.4, we have

$$\begin{aligned}
A_\infty &= \frac{1}{2}[(L^{**})^c]_\infty + \sum_{0 < s} \log(1 + \Delta L_s^{**}) - \frac{\Delta L_s^{**}}{1 + \Delta L_s^{**}} \\
&= \sum_{i=1}^d \int_u^t \log(1 + (H^{**})_s^i) - \frac{(H^{**})_s^i}{1 + (H^{**})_s^i} dN_s^i \leq \sum_{i=1}^d \int_u^t \log \frac{\mu_s^i}{(\mu^-)_s^i} dN_s^i \\
(A.18) \quad &= \sum_{i=1}^d \int_u^t 1_{(\mu_s^i \geq \lambda_s^i)} \log \frac{\mu_s^i}{\lambda_s^i} dN_s^i = \sum_{i=1}^d \int_u^t \log_+ \gamma_s^i dN_s^i.
\end{aligned}$$

Also, since  $-1 \leq H^* \leq 0$ , we find

$$\begin{aligned}
\mathcal{E}(L^*)_\infty &= \exp \left( \sum_{i=1}^d - \int_u^t (H^*)_s^i \lambda_s^i ds + \int_u^t \log(1 + (H^*)_s^i) dN_s^i \right) \\
(A.19) \quad &\leq \exp \left( \sum_{i=1}^d \int_u^t \lambda_s^i ds \right),
\end{aligned}$$

which leads to

$$\begin{aligned}
E^Q \exp(A_\infty) &= E \mathcal{E}(L^*)_\infty \exp(A_\infty) \\
(A.20) \quad &\leq E \exp \left( \sum_{i=1}^d \int_u^t \lambda_s^i ds + \int_u^t \log_+ \gamma_s^i dN_s^i \right),
\end{aligned}$$

which is finite by assumption. Theorem 3.4 then shows that  $L^{**}$  is a uniformly integrable martingale under  $Q$ , so  $E^Q \mathcal{E}(L^{**})_\infty = 1$ , from which we conclude  $E^P \mathcal{E}(L)_\infty = 1$ . Thus,  $\mathcal{E}(L)$  is a uniformly integrable martingale, in particular a martingale. This completes the proof.  $\square$

**Proof of Corollary 2.3.** First note that for  $x \geq 0$ , it holds that

$$(A.21) \quad x \log x - (x - 1) \leq 1 + x \log x \leq 1 + x \log_+ x.$$

Therefore, as  $\lambda = 1$ , the first moment condition of Theorem 2.2 reduces to the first moment condition in the statement of the corollary. Furthermore, because of  $\lambda = 1$ , we obtain

$$(A.22) \quad E \exp \left( \sum_{i=1}^d \int_u^t \lambda_s^i ds + \int_u^t \log_+ \gamma_s^i dN_s^i \right) = e^{d(t-u)} E \exp \left( \sum_{i=1}^d \int_u^t \log_+ \gamma_s^i dN_s^i \right),$$

and the result for the second moment condition of the corollary follows. This completes the proof.  $\square$

## A.2. Proofs for Section 4.



**Proof of Example 4.1.** By our assumptions,  $\gamma_t = \alpha\lambda_t^{-1} + \beta \leq \alpha\delta^{-1} + \beta$ . Using that  $x \log x - (x - 1) \leq 1 + x \log x \leq 1 + x \log_+ x$  for any  $x \geq 0$ , we then obtain

$$(A.23) \quad \int_u^t (\gamma_s^i \log \gamma_s^i - (\gamma_s^i - 1)) \lambda_s \, ds \leq (1 + (\alpha\delta^{-1} + \beta) \log_+(\alpha\delta^{-1} + \beta)) \int_u^t \lambda_s \, ds,$$

so the first criterion of Theorem 2.2 yields the result.  $\square$

**Proof of Example 4.2.** Let  $\varepsilon > 0$  be given such that  $\exp(\varepsilon \langle H \cdot M \rangle_t)$  is integrable for all  $t \geq 0$ . Pick  $M > 0$  so large that  $x \log_+ x \leq \varepsilon(x - 1)^2$  for  $x \geq M$ . Defining  $C = \sup_{-1 \leq x \leq M} x \log_+ x$ , we then obtain for  $0 \leq u \leq t$  that

$$(A.24) \quad \begin{aligned} E \exp \left( \sum_{i=1}^d \int_u^t \mu_s^i \log_+ \mu_s^i \, ds \right) &\leq E \exp \left( \sum_{i=1}^d \int_0^t C + 1_{(\mu_s^i \geq M)} \varepsilon (\mu_s^i - 1)^2 \, ds \right) \\ &\leq \exp(dtC) E \exp \left( \varepsilon \sum_{i=1}^d \int_0^t (H_s^i)^2 \, ds \right). \end{aligned}$$

As  $N$  has no common jumps, however, we have

$$(A.25) \quad [H \cdot M]_t = \left[ \sum_{i=1}^d H^i \cdot M^i \right]_t = \sum_{i=1}^d (H^i)^2 \cdot [M^i]_t = \sum_{i=1}^d \int_0^t (H_s^i)^2 \, dN_s^i,$$

and therefore,  $\langle H \cdot M \rangle_t = \sum_{i=1}^d \int_0^t (H_s^i)^2 \, ds$ . All in all, we conclude

$$(A.26) \quad E \exp \left( \sum_{i=1}^d \int_u^t \mu_s^i \log_+ \mu_s^i \, ds \right) \leq E \exp(dtC) \exp(\varepsilon \langle H \cdot M \rangle_t) < \infty,$$

and the result follows by Corollary 2.3.  $\square$

**Lemma A.4.** *There is  $C$  such that  $\frac{1}{n!} \leq C \exp(n - n \log n)$  for all  $n \geq 0$ , using the convention that  $0 \log 0 = 0$  and  $0! = 1$ .*

*Proof.* By (6.11.2) of (Zwillinger, 2003), it holds that  $\lim_n n! (\sqrt{2\pi n} (n/e)^n)^{-1} = 1$ . Fix  $\varepsilon$  in  $(0, 1)$ . For  $n$  large enough,  $n! \geq (1 - \varepsilon) \sqrt{2\pi n} (n/e)^n = (1 - \varepsilon) \sqrt{2\pi} \exp((n + \frac{1}{2}) \log n - n)$ . This implies  $n! \geq C^{-1} \exp((n + \frac{1}{2}) \log n - n)$  for a positive constant  $C$  and all  $n \geq 0$ . From this, we conclude  $\frac{1}{n!} \leq C \exp(n - (n + \frac{1}{2}) \log n) \leq C \exp(n - n \log n)$  for all  $n \geq 0$ .  $\square$

**Lemma A.5.** *Let  $Z$  be Poisson distributed with parameter  $\mu$ . Then  $\exp(\varepsilon Z \log Z)$  is integrable whenever  $0 \leq \varepsilon < 1$ .*

*Proof.* We find, using Lemma A.4,

$$\begin{aligned}
 E \exp(\varepsilon Z \log Z) &= \sum_{n=0}^{\infty} \exp(\varepsilon n \log n) \frac{\mu^n}{n!} \exp(-\mu) \\
 &\leq C \sum_{n=0}^{\infty} \exp(\varepsilon n \log n - n \log \mu + n - n \log n) \\
 (A.27) \quad &= C \sum_{n=0}^{\infty} (\exp(1 - \log \mu) n^{\varepsilon-1})^n.
 \end{aligned}$$

As  $\varepsilon - 1 < 0$ ,  $|\exp(1 - \log \mu) n^{\varepsilon-1}| < 1$  for  $n$  large enough. Therefore, the above series is convergent by comparison to a geometric series.  $\square$

**Proof of Example 4.3 using the first moment condition of Corollary 2.3.** As  $x \log_+ x$  is increasing in  $x$ , it suffices to consider the case where  $\alpha > 1$  and  $\beta > 0$ , such that  $\mu$  is positive. Fix  $\varepsilon > 0$ , and let  $0 \leq u \leq t$  with  $|t - u| \leq \varepsilon$ . We then obtain, with  $N_t^S = \sum_{j=1}^d N_t^j$ ,

$$\begin{aligned}
 \exp \left( \sum_{i=1}^d \int_u^t \mu_s^i \log_+ \mu_s^i ds \right) &\leq \exp \left( \sum_{i=1}^d \int_u^t (\alpha + \beta N_{s-}^S) \log(\alpha + \beta N_{s-}^S) ds \right) \\
 (A.28) \quad &\leq \exp(\varepsilon d(\alpha + \beta N_t^S) \log(\alpha + \beta N_t^S))
 \end{aligned}$$

Now, for  $k$  large, we have  $\alpha + \beta k \leq 2\beta k$  and  $\log(2\beta k) \leq 2 \log k$ , so for  $k$  large enough, we find  $\varepsilon d(\alpha + \beta k) \log(\alpha + \beta k) \leq 2\varepsilon d\beta k \log(2\beta k) \leq 4\varepsilon d\beta k \log k$ . From this we conclude that  $\exp(\sum_{i=1}^d \int_u^t \mu_s^i \log \mu_s^i ds)$  is integrable if only  $\exp(4\varepsilon d\beta N_t^S \log N_t^S)$  is integrable.  $N_t^S$  is Poisson distributed with parameter  $dt$ , so by choosing  $\varepsilon$  with  $4\varepsilon d\beta < 1$ , we obtain the desired integrability using Lemma A.5. Corollary 2.3 now yields the result.  $\square$

**Proof of Example 4.3 using the second moment condition of Corollary 2.3.**

Again, it suffices to consider  $\alpha > 1$  and  $\beta > 0$ . We fix  $\varepsilon > 0$  and consider  $0 \leq u \leq t$  satisfying  $|t - u| \leq \varepsilon$ . We put  $N_t^S = \sum_{j=1}^d N_t^j$  and define a mapping  $\varphi : \mathbb{N}_0 \rightarrow \mathbb{R}$  by  $\varphi(n) = E \exp(\int_0^{t-u} \log \beta(n + N_s^S) dN_s^S)$ . Let  $m \in \mathbb{N}$  such that  $\alpha \leq \beta m$ , we then obtain  $\alpha + \beta x \leq \beta(m + x)$ . As  $N^S$  is a Poisson process of rate  $d$ , we find that conditionally on  $N_u^S$ , the processes  $s \mapsto N_{s+u}^S - N_u^S$  and  $s \mapsto N_s^S$  have the same distribution, and we obtain by conditioning on  $N_u^S$  that

$$\begin{aligned}
 (A.29) \quad E \exp \left( \sum_{i=1}^d \int_u^t \log \mu_s^i dN_s^i \right) &\leq E \exp \left( \sum_{i=1}^d \int_u^t \log \beta \left( m + \sum_{j=1}^d N_s^j \right) dN_s^i \right) \\
 &= E \exp \left( \int_u^t \log \beta(m + N_s^S) dN_s^S \right) = E \exp \left( \int_u^t \log \beta(m + N_u^S + N_s^S - N_u^S) dN_s^S \right) \\
 &= E \exp \left( \int_0^{t-u} \log \beta(m + N_u^S + N_{s+u}^S - N_u^S) d(N_{s+u}^S - N_u^S) \right) = E \varphi(m + N_u^S).
 \end{aligned}$$

Now note that

$$\begin{aligned}
\varphi(n) &= E \exp \left( \int_0^{t-u} \log \beta(n + N_s^S) dN_s^S \right) = E \exp \left( \sum_{k=1}^{N_{t-u}^S} \log \beta(n + k) \right) \\
&= \sum_{p=0}^{\infty} \exp \left( \sum_{k=1}^p \log \beta(n + k) \right) \frac{((t-u)d)^p}{p!} \exp(-(t-u)d) \\
&= \exp(-(t-u)d) \sum_{p=0}^{\infty} \beta^p \left( \prod_{k=1}^p (n+k) \right) \frac{((t-u)d)^p}{p!} \\
(A.30) \quad &= \exp(-(t-u)d) \sum_{p=0}^{\infty} (\beta(t-u)d)^p \frac{(n+p)!}{n!p!}.
\end{aligned}$$

Whenever  $|x| < 1$ , we have  $\sum_{p=0}^{\infty} x^p \frac{(n+p)!}{n!p!} = (1-x)^{-(n+1)}$  by formula (15.1.8) of (Abramowitz and Stegun, 1964), and we therefore conclude, whenever  $\beta(t-u)d < 1$ , that

$$(A.31) \quad \varphi(n) = \frac{\exp(-(t-u)d)}{(1-\beta(t-u)d)^{n+1}}.$$

Therefore, in this case,

$$\begin{aligned}
E \exp \left( \sum_{i=1}^d \int_u^t \log \mu_s^i dN_s^i \right) &\leq E \varphi(m + N_u^S) = E \frac{\exp(-(t-u)d)}{(1-\beta(t-u)d)^{m+N_u^S+1}} \\
&= \exp(-td) \sum_{p=0}^{\infty} (1-\beta(t-u)d)^{-(p+m+1)} \frac{(ud)^p}{p!} \\
&= \frac{1}{(1-\beta(t-u)d)^{m+1}} \exp(-td) \sum_{p=0}^{\infty} \left( \frac{ud}{1-\beta(t-u)d} \right)^p \frac{1}{p!} \\
(A.32) \quad &= \frac{1}{(1-\beta(t-u)d)^{m+1}} \exp \left( -td + \frac{ud}{1-\beta(t-u)d} \right).
\end{aligned}$$

We conclude that the second moment condition of Corollary 2.3 yields the result, using  $\varepsilon$  such that  $\beta\varepsilon d < 1$ .  $\square$

For the diffusion examples we need two lemmas.

**Lemma A.6.** *Consider three mappings  $A : \mathbb{N}_0^d \times \mathbb{R}_+^d \rightarrow \mathbb{R}^d$ ,  $B : \mathbb{N}_0^d \times \mathbb{R}_+^d \rightarrow \mathbb{M}(d, d)$  and  $\sigma : \mathbb{N}_0^d \times \mathbb{R}_+^d \rightarrow \mathbb{M}(d, d)$  such that  $A(\eta, \cdot)$ ,  $B(\eta, \cdot)$  and  $\sigma(\eta, \cdot)$  are bounded and continuous for  $\eta \in \mathbb{N}_0^d$ . Let  $W$  be a  $d$ -dimensional  $(\mathcal{F}_t)$  Brownian motion. Let  $T_n^i$  be the  $n$ 'th event time for  $N^i$  and let  $Z_t^i = t - T_{N_t^i}^i$ . The stochastic differential equation*

$$(A.33) \quad dX_t = (A(N_t, Z_t) + B(N_t, Z_t)X_t) dt + \sigma(N_t, Z_t) dW_t$$

is exact, in the sense that for any initial value, it has a pathwise unique solution. Defining  $C_t = \exp(-\int_0^t B(N_s, Z_s) ds)$ , the solution is

$$(A.34) \quad X_t = C_t^{-1} \left( X_0 + \int_0^t C_s A(N_s, Z_s) ds + \int_0^t C_s \sigma(N_s, Z_s) dW_s \right).$$

*Proof.* Let  $\tilde{A}_s = A(N_s, Z_s)$ , and define  $\tilde{B}$  and  $\tilde{\sigma}$  analogously. Note that as  $N$  and  $Z$  are adapted,  $\tilde{A}$  is adapted as well, since  $A(\eta, \cdot)$  is continuous and so Borel measurable for all  $\eta \in \mathbb{N}_0^d$ . As the process is also right-continuous and locally bounded, all integrals are well-defined, and similarly for  $\tilde{B}$  and  $\tilde{\sigma}$ . Let  $X_0$  be some initial value. Assume that  $X$  is a solution to the stochastic differential equation. Note that each entry of  $C_t$  is differentiable as a function of  $t$ , and  $\frac{d}{dt} C_t^{ij} = (-\tilde{B}_t C_t)^{ij}$ . The integration-by-parts formula yields

$$(A.35) \quad \begin{aligned} (C_t X_t)_i &= \sum_{j=1}^d C_t^{ij} X_t^j = X_0^i + \sum_{j=1}^d \int_0^t C_s^{ij} dX_s^j + \int_0^t X_s^j dC_s^{ij} \\ &= X_0^i + \sum_{j=1}^d \int_0^t C_s^{ij} (\tilde{A}_s^j + \sum_{k=1}^d \tilde{B}_s^{jk} X_s^k) ds + \int_0^t C_s^{ij} \sum_{k=1}^d \tilde{\sigma}_s^{jk} dW_s^k + \int_0^t X_s^j (-\tilde{B}_s C_s)^{ij} ds \end{aligned}$$

Relabeling the indicies, we find that

$$(A.36) \quad \sum_{j=1}^d \int_0^t C_s^{ij} \sum_{k=1}^d \tilde{B}_s^{jk} X_s^k ds = \sum_{j=1}^d \sum_{k=1}^d \int_0^t C_s^{ik} \tilde{B}_s^{kj} X_s^j ds = \sum_{j=1}^d \int_0^t X_s^j (C_s \tilde{B}_s)^{ij} ds,$$

and so  $(C_t X_t)_i = X_0^i + \sum_{j=1}^d \int_0^t C_s^{ij} \tilde{A}_s^j ds + \int_0^t C_s^{ij} \sum_{k=1}^d \tilde{\sigma}_s^{jk} dW_s^k$ . With the usual matrix notation for integrals, this means that  $C_t X_t = X_0^i + \int_0^t C_s \tilde{A}_s ds + \int_0^t C_s \tilde{\sigma}_s dW_s$ , and so we obtain

$$(A.37) \quad X_t = C_t^{-1} \left( X_0 + \int_0^t C_s A(N_s, Z_s) ds + \int_0^t C_s \sigma(N_s, Z_s) dW_s \right).$$

This proves pathwise uniqueness. Applying the integration-by-parts formula to the above shows that the proposed solution in fact is a solution. This proves existence.  $\square$

**Lemma A.7.** *Let  $X$  be a  $d$ -dimensional normally distributed variable with mean  $\xi$  and positive definite variance  $\Sigma$ . Let  $c > 0$  and  $0 < \varepsilon < 1$ . Then  $\exp(c\|X\|_2^{1+\varepsilon})$  is integrable. Furthermore, defining  $a(c, \varepsilon) = 2^{1+\varepsilon}c$  and  $b(c, \varepsilon) = 16^{(1+\varepsilon)/(1-\varepsilon)}c^{2/(1-\varepsilon)}$ , it holds that*

$$(A.38) \quad E \exp(c\|X\|_2^{1+\varepsilon}) \leq k_d \exp(a(c, \varepsilon)\|\xi\|^{1+\varepsilon}) \exp\left(b(c, \varepsilon)\|\Sigma\|_2^{\frac{1+\varepsilon}{1-\varepsilon}}\right),$$

where  $k_d = A_d m_{d-1}(\sqrt{2}\sqrt{\pi^{d-1}})^{-1}$ ,  $A_d$  is the area of the unit sphere in  $d$  dimensions and  $m_d$  is the  $d$ 'th absolute moment of the standard normal distribution.

*Proof.* By (Lancaster and Tismenetsky, 1985), p. 181,  $\Sigma$  has a unique symmetric positive definite square root  $\Sigma^{1/2}$  such that  $\Sigma = (\Sigma^{1/2})^2$ . Furthermore, with  $Y = \Sigma^{-1/2}(X - \xi)$ , it holds that  $X = \xi + \Sigma^{\frac{1}{2}}Y$ , where  $Y$  is  $d$ -dimensionally standard normally distributed.

Note that for  $x, y \geq 0$ , it holds that  $(x + y)^{1+\varepsilon} \leq (2 \max\{x, y\})^{1+\varepsilon} \leq 2^{1+\varepsilon}(x^{1+\varepsilon} + y^{1+\varepsilon})$ . Also applying  $\|\Sigma^{1/2}\|_2 = \sqrt{\|\Sigma\|_2}$ , where  $\|\cdot\|_2$  is the operator norm induced by the Euclidean norm, we get

$$\begin{aligned}
 E \exp(c\|X\|_2^{1+\varepsilon}) &= E \exp(c\|\xi + \Sigma^{\frac{1}{2}}Y\|_2^{1+\varepsilon}) \leq E \exp(c(\|\xi\|_2 + \|\Sigma^{\frac{1}{2}}Y\|_2)^{1+\varepsilon}) \\
 &\leq E \exp(c2^{1+\varepsilon}(\|\xi\|_2^{1+\varepsilon} + \|\Sigma^{\frac{1}{2}}Y\|_2^{1+\varepsilon})) \\
 (A.39) \quad &\leq \exp(c2^{1+\varepsilon}\|\xi\|_2^{1+\varepsilon}) E \exp\left(c2^{1+\varepsilon}\|\Sigma\|_2^{(1+\varepsilon)/2}\|Y\|_2^{1+\varepsilon}\right).
 \end{aligned}$$

Switching to polar coordinates (see (Rudin, 1970), page 149) we obtain, with  $A_d$  denoting the area of the unit sphere in  $d$  dimensions and  $C = c2^{1+\varepsilon}\|\Sigma\|_2^{(1+\varepsilon)/2}$ ,

$$\begin{aligned}
 &E \exp\left(c2^{1+\varepsilon}\|\Sigma\|_2^{(1+\varepsilon)/2}\|Y\|_2^{1+\varepsilon}\right) \\
 &= \int_{\mathbb{R}^d} \exp(C\|x\|_2^{1+\varepsilon}) \frac{1}{\sqrt{(2\pi)^d}} \exp\left(-\frac{1}{2}\|x\|_2^2\right) dx \\
 (A.40) \quad &= \frac{A_d}{\sqrt{(2\pi)^{d-1}}} \int_0^\infty \exp(Cr^{1+\varepsilon}) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}r^2\right) r^{d-1} dr.
 \end{aligned}$$

Using a change of variables, we obtain the bound

$$\begin{aligned}
 &\int_0^\infty \exp(Cr^{1+\varepsilon}) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}r^2\right) r^{d-1} dr \\
 &= \int_0^\infty \exp\left(Cr^{1+\varepsilon} - \frac{1}{4}r^2\right) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{4}r^2\right) r^{d-1} dr \\
 &\leq \int_0^\infty r^{d-1} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{4}r^2\right) dr \sup_{s \geq 0} \exp\left(Cs^{1+\varepsilon} - \frac{1}{4}s^2\right) \\
 (A.41) \quad &= 2^{d/2} \int_0^\infty r^{d-1} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}r^2\right) dr \sup_{s \geq 0} \exp\left(Cs^{1+\varepsilon} - \frac{1}{4}s^2\right).
 \end{aligned}$$

With  $m_d$  denoting the  $d$ 'th absolute moment of the standard normal distribution, we have  $\int_0^\infty r^{d-1} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}r^2\right) dr = \frac{1}{2}m_{d-1}$ . Also, defining  $\phi(r) = Cr^{1+\varepsilon} - \frac{1}{4}r^2$  for  $r \geq 0$ ,  $\phi$  has a global maximum at  $r^* = (2C(1+\varepsilon))^{1/(1-\varepsilon)}$  which satisfies

$$\begin{aligned}
 \phi(r^*) &= C \left((2C(1+\varepsilon))^{1/(1-\varepsilon)}\right)^{1+\varepsilon} - \frac{1}{4} \left((2C(1+\varepsilon))^{1/(1-\varepsilon)}\right)^2 \\
 (A.42) \quad &\leq 2^{\frac{1+\varepsilon}{1-\varepsilon}} C^{\frac{2}{1-\varepsilon}} (1+\varepsilon)^{\frac{1+\varepsilon}{1-\varepsilon}} \leq 4^{\frac{1+\varepsilon}{1-\varepsilon}} C^{\frac{2}{1-\varepsilon}}.
 \end{aligned}$$

As the exponential mapping is increasing, this allows us to conclude

$$(A.43) \quad \int_0^\infty \exp(Cr^{1+\varepsilon}) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}r^2\right) r^{d-1} dr \leq 2^{d/2-1} m_{d-1} \exp\left(4^{\frac{1+\varepsilon}{1-\varepsilon}} C^{\frac{2}{1-\varepsilon}}\right).$$

Recalling our definition of  $C$ , we have

$$(A.44) \quad 4^{\frac{1+\varepsilon}{1-\varepsilon}} C^{\frac{2}{1-\varepsilon}} = 4^{\frac{1+\varepsilon}{1-\varepsilon}} \left(c2^{1+\varepsilon}\|\Sigma\|_2^{(1+\varepsilon)/2}\right)^{\frac{2}{1-\varepsilon}} = 16^{\frac{1+\varepsilon}{1-\varepsilon}} c^{\frac{2}{1-\varepsilon}} \|\Sigma\|_2^{\frac{1+\varepsilon}{1-\varepsilon}}.$$

Therefore, defining  $a(c, \varepsilon) = 2^{1+\varepsilon}c$  and  $b(c, \varepsilon) = 16^{(1+\varepsilon)/(1-\varepsilon)}c^{2/(1-\varepsilon)}$ , we finally conclude

$$(A.45) \quad E \exp(c\|X\|_2^{1+\varepsilon}) \leq \frac{A_d m_{d-1}}{\sqrt{2}\sqrt{\pi^{d-1}}} \exp(a(c, \varepsilon)\|\xi\|_2^{1+\varepsilon}) \exp\left(b(c, \varepsilon)\|\Sigma\|_2^{\frac{1+\varepsilon}{1-\varepsilon}}\right),$$

which proves the lemma.  $\square$

**Proof of Example 4.4.** We need to show that the first criterion of Corollary 2.3 is applicable. We may assume without loss of generality that  $0 < \delta < 1$ . It suffices to prove that for any  $t > 0$ ,  $E \exp(\sum_{i=1}^d \int_0^t \mu_s^i \log_+ \mu_s^i ds)$  is finite. Fix  $t > 0$ . By Jensen's inequality, we find

$$(A.46) \quad E \exp \left( \sum_{i=1}^d \int_0^t \mu_s^i \log_+ \mu_s^i ds \right) \leq \frac{1}{t} \int_0^t E \exp \left( t \sum_{i=1}^d \mu_s^i \log_+ \mu_s^i \right) ds.$$

We wish to bound the expectation inside the integral by an expression depending continuously on  $s$ . Recall that we have assumed that  $\phi$  is Lipschitz. As all norms on finite-dimensional vector spaces are equivalent, we find in particular that there exists  $\gamma > 0$  such that  $\|\phi(x)\|_\infty \leq \gamma \|x\|_2$ . As  $\phi(x) \in \mathbb{R}_+^d$ , we obtain  $\phi^i(x) \leq \gamma \|x\|_2$  for all  $i \leq d$ , and so  $E \exp(t \sum_{i=1}^d \mu_s^i \log_+ \mu_s^i) \leq E \exp(td\gamma \|X_s\|_2 \log_+ \gamma \|X_s\|_2)$ . Next, let  $0 < \zeta < 1$ . It holds for all  $x \geq 0$  that  $\log_+ x \leq \zeta^{-1} x^\zeta$ . Therefore, defining  $\rho = td\gamma^{1+\zeta} \zeta^{-1}$ , we conclude

$$(A.47) \quad E \exp \left( t \sum_{i=1}^d \mu_s^i \log_+ \mu_s^i \right) \leq E \exp \left( \rho \|X_s\|_2^{\zeta+1} \right).$$

We will calculate this expectation by conditioning on  $N$ . Let  $\eta$  denote a point process path, and let  $(\tau_n)$  denote the event times of  $\eta$ . By the explicit representation in Lemma A.6 as well as the results on pathwise stochastic integration in (Karandikar, 1995), it holds that conditionally on  $N = \eta$ ,  $X_s$  has the same distribution as  $Y_s^\eta$ , where

$$(A.48) \quad Y_s^\eta = C_s^{-1} \left( x_0 + \int_0^s C_v A(\eta_v, v - \tau_{\eta_v}) dv + \int_0^s C_v \sigma(\eta_v, v - \tau_{\eta_v}) dW_v \right),$$

which is a normal distribution with mean  $\xi_s^\eta$  and variance  $\Sigma_s^\eta$ , where

$$(A.49) \quad \xi_s^\eta = C_s^{-1} \left( x_0 + \int_0^s C_v A(\eta_v, v - \tau_{\eta_v}) dv \right)$$

$$(A.50) \quad \Sigma_s^\eta = C_s^{-1} \int_0^s (C_v \sigma(\eta_v, v - \tau_{\eta_v}))^t (C_v \sigma(\eta_v, v - \tau_{\eta_v})) ds (C^{-1})_s^t,$$

and where  $C_s = \exp(-\int_0^s B(\eta_v, v - \tau_{\eta_v}) dv)$ . With  $\|\cdot\|_2$  denoting the matrix operator norm induced by the Euclidean norm, Lemma A.7 yields

$$(A.51) \quad \begin{aligned} E \exp(\rho \|X_s\|_2^{1+\zeta}) &= \int E \left( \exp(\rho \|X_s\|_2^{1+\zeta}) \middle| N = \eta \right) dN(P)(\eta) \\ &= \int E \exp(\rho \|Y_s^\eta\|_2^{1+\zeta}) dN(P)(\eta) \leq k_d E \exp(a(\rho, \zeta) \|\xi_s^\eta\|_2^{1+\zeta}) \exp \left( b(\rho, \zeta) \|\Sigma_s^\eta\|_2^{\frac{1+\zeta}{1-\zeta}} \right). \end{aligned}$$

Next, we consider bounds for  $\|\xi_s^\eta\|_2$  and  $\|\Sigma_s^\eta\|_2$ . We begin by noting that we always have  $\|C_s\|_2 \leq \exp(\int_0^s \|B(\eta_v, v - \tau_{\eta_v})\|_2 dv) \leq \exp(sc_B)$ , where we have applied standard norm inequalities, see Theorem 10.10 of (Higham, 2008) and Lemma 1.4 of (Ethier and Kurtz,

1986), and similarly,  $\|C_s^{-1}\|_2 \leq \exp(sc_B)$ . Therefore, recalling that  $0 < \delta < 1$  so that  $x \mapsto x^{1-\delta}$  is increasing,

$$\begin{aligned}
\|\xi_s^\eta\|_2 &\leq \exp(sc_B) \left( \|x_0\|_2 + \int_0^s \|C_v\|_2 \|A(\eta_v, v - \tau_{\eta_v})\|_2 dv \right) \\
&\leq \exp(sc_B) \left( \|x_0\|_2 + \exp(sc_B) \int_0^s \|A(\eta_v)\|_2 dv \right) \\
(A.52) \quad &\leq \exp(sc_B) (\|x_0\|_2 + sc_A \exp(sc_B) \|\eta_s\|_1^{1-\delta}).
\end{aligned}$$

Similarly, we obtain

$$\begin{aligned}
\|\Sigma_s^\eta\|_2 &\leq \exp(2sc_B) \left\| \int_0^s (C_v \sigma(\eta_v, v - \tau_{\eta_v}))^t (C_v \sigma(\eta_v, v - \tau_{\eta_v})) ds \right\|_2 \\
(A.53) \quad &\leq \exp(2sc_B) \int_0^s \exp(2vc_B) \|\sigma(\eta_v, v - \tau_{\eta_v})\|_2^2 dv \leq s \exp(4sc_B) c_\sigma^2 \|\eta_s\|_1^{1-\delta}.
\end{aligned}$$

In particular, for appropriate continuous functions  $a_\xi$ ,  $b_\xi$  and  $b_\Sigma$  from  $\mathbb{R}_+$  to  $\mathbb{R}$ , depending on  $\zeta$ , we obtain the two bounds

$$(A.54) \quad \|\xi_s^\eta\|_2^{1+\zeta} \leq a_\xi(s) + b_\xi(s) \|\eta_s\|_1^{(1-\delta)(1+\zeta)}$$

$$(A.55) \quad \|\Sigma_s^\eta\|_2^{\frac{1+\zeta}{1-\zeta}} \leq b_\Sigma(s) \|\eta_s\|_1^{(1-\delta)\frac{1+\zeta}{1-\zeta}}.$$

We then conclude

$$\begin{aligned}
&E \exp(\rho \|X_s\|_2^{1+\zeta}) \\
&\leq k_d E \exp \left( a(\rho, \zeta) \left( a_\xi(s) + b_\xi(s) \|N_s\|_1^{(1-\delta)(1+\zeta)} \right) + b(\rho, \zeta) b_\Sigma(s) \|N_s\|_1^{(1-\delta)\frac{1+\zeta}{1-\zeta}} \right) \\
(A.56) \quad &\leq k_d \exp(a(\rho, \zeta) a_\xi(s)) E \exp \left( (a(\rho, \zeta) b_\xi(s) + b(\rho, \zeta) b_\Sigma(s)) \|N_s\|_1^{(1-\delta)\frac{1+\zeta}{1-\zeta}} \right).
\end{aligned}$$

The above depends on given constants  $\delta$ ,  $c_A$ ,  $c_B$  and  $c_\sigma$ , as well as the constant  $\zeta$  which we may choose arbitrarily in the open interval between zero and one. We now choose  $\zeta$  so small in  $(0, 1)$  that  $(1-\delta)(1+\zeta)(1-\zeta)^{-1} \leq 1$ . We then also obtain  $(1-\delta)(1+\zeta) \leq 1$ . Recalling that for any Poisson distributed variable  $Z$  with intensity  $\lambda$  and any  $c \in \mathbb{R}$ , it holds that  $E \exp(cZ) = \exp((\exp(c) - 1)\lambda)$ , we may then conclude

$$\begin{aligned}
&E \exp(\rho \|X_s\|_2^{1+\zeta}) \\
&\leq k_d \exp(a(\rho, \zeta) a_\xi(s)) E \exp((a(\rho, \zeta) b_\xi(s) + b(\rho, \zeta) b_\Sigma(s)) \|N_s\|_1) \\
(A.57) \quad &= k_d \exp(a(\rho, \zeta) a_\xi(s)) \exp((\exp(a(\rho, \zeta) b_\xi(s) + b(\rho, \zeta) b_\Sigma(s)) - 1) ds).
\end{aligned}$$

All in all, we may now define, for  $0 \leq s \leq t$ ,

$$(A.58) \quad \varphi(s) = k_d \exp(a(\rho, \zeta) a_\xi(s)) \exp((a(\rho, \zeta) \exp(b_\xi(s) + b(\rho, \zeta) b_\Sigma(s)) - 1) ds),$$

and obtain  $E \exp(t \sum_{i=1}^d \mu_s^i \log_+ \mu_s^i) \leq \varphi(s)$  for all such  $s$ . The functions  $a_\xi$ ,  $b_\xi$  and  $b_\Sigma$  depends continuously on  $s$ . Therefore,  $\varphi$  is a continuous function of  $s$ . In particular, the integral of  $\varphi$  over  $[0, t]$  is finite. Recalling our first estimates, this leads us to conclude that for any  $t \geq 0$ , it holds that  $E \exp(\sum_{i=1}^d \int_0^t \mu_s^i \log_+ \mu_s^i ds)$  is finite, and so the first integrability criterion of Corollary 2.3 is satisfied.  $\square$

**Proof of Example 4.5.** We want to show that the second moment condition of Corollary 2.3 is applicable. To this end, we first construct an explicit solution to the stochastic differential equation defining  $X$ . With  $T_n$  denoting the  $n$ 'th event time for  $N$ , define the process  $W^n$  by  $W_t^n = W_{T_n+t} - W_{T_n}$  and define  $\mathcal{F}_t^n = \mathcal{F}_{T_n+t}$ . By Theorem I.12.1 of (Rogers and Williams, 2000a),  $W^n$  is independent of  $\mathcal{F}_{T_n}$  and has the distribution of a Brownian motion. Again using Theorem I.12.1 of (Rogers and Williams, 2000a) with the stopping time  $T_n + s$ , we have for  $0 \leq s \leq t$  that

$$\begin{aligned}
 E(W_t^n | \mathcal{F}_s^n) &= E(W_{T_n+t} - W_{T_n} | \mathcal{F}_{T_n+s}) \\
 &= E(W_{T_n+t} - W_{T_n+s} | \mathcal{F}_{T_n+s}) + W_{T_n+s} - W_{T_n} \\
 (A.59) \quad &= W_{T_n+s} - W_{T_n} = W_s^n,
 \end{aligned}$$

and Lévy's characterisation Theorem for Brownian motion relative to a filtration, see (Rogers and Williams, 2000b), Theorem IV.33.1, shows that  $W^n$  is an  $(\mathcal{F}_t^n)$ -Brownian motion. We may then use the Itô existence and uniqueness theorem, see Theorem 11.2 of (Rogers and Williams, 2000b), concluding that on the same probability space that carries the Poisson process  $N$ , the Brownian motion  $W$  and in particular the  $(\mathcal{F}_t^n)$ -Brownian motion  $W^n$ , there exist unique processes  $X^n$  satisfying the stochastic differential equations  $dX_t^n = a_n + b_n X_t^n dt + \sigma dW_t^n$  with constant initial values  $\xi_n$ . Whenever  $T_n \leq t < T_{n+1}$ , we then have

$$\begin{aligned}
 X_{t-T_n}^n &= \xi_n + \int_0^{t-T_n} a_n + b_n X_s^n ds + \int_0^{t-T_n} \sigma dW_s^n \\
 (A.60) \quad &= \xi_n + \int_{T_n}^t a_n + b_n X_{s-T_n}^n ds + \int_{T_n}^t \sigma dW_s.
 \end{aligned}$$

The process  $\sum_{n=0}^{\infty} X_{t-T_n}^n 1_{[T_n, T_{n+1})}(t)$  thus satisfies the same stochastic differential equation as  $X$ . By pathwise uniqueness for each  $X^n$ , we find  $X_t = \sum_{n=0}^{\infty} X_{t-T_n}^n 1_{[T_n, T_{n+1})}(t)$ .

The above deliberations yield an explicit representation for the stochastic differential equation defining the intensity. Next, we check that the second moment condition of Corollary 2.3 is applicable. With  $S_k = T_k - T_{k-1}$  denoting the sequence of interarrival times, we then obtain for the moment condition to be investigated that

$$\begin{aligned}
 E \exp \left( \int_u^t \log_+ |X_{s-}| dN_s \right) &\leq E \exp \left( \int_u^t \log(1 + |X_{s-}|) dN_s \right) \\
 &= E \prod_{k=N_u+1}^{N_t} (1 + |X_{T_k-}|) = E \prod_{k=N_u+1}^{N_t} (1 + |X_{T_k-T_{k-1}}^{k-1}|) \\
 (A.61) \quad &= E \prod_{k=N_u+1}^{N_t} (1 + |X_{S_k}^{k-1}|).
 \end{aligned}$$

In order to obtain the finiteness of this expression, we wish to condition on  $N$ . Given a counting process trajectory  $\eta$ , we refer to the event times of  $\eta$  by  $(\tau_n)$ ,  $\tau_0 = 0$ , and we



let  $(s_n)$  be the corresponding interarrival times,  $s_n = \tau_n - \tau_{n-1}$ . We then have

$$(A.62) \quad E \prod_{k=N_u+1}^{N_t} (1 + |X_{S_k}^{k-1}|) = \int E \left( \prod_{k=\eta_u+1}^{\eta_t} (1 + |X_{s_k}^{k-1}|) \middle| N = \eta \right) dN(P)(\eta),$$

Next, we argue that given  $N$ , the variables  $(X_{s_k}^{k-1})_{k \geq 1}$  are mutually independent, in the sense that it  $N(P)$  almost surely holds that the conditional distribution of the variables  $(X_{s_k}^{k-1})_{k \geq 1}$  given  $N = \eta$  is the product measure of each of the marginal conditional distributions.

To this end, we begin by arguing that each  $X_{s_k}^{k-1}$  is a transformation of  $(W^{k-1})^{s_k}$ . By the definition of  $X^k$ ,  $X_{s_k}^{k-1} = \xi_{k-1} + \int_0^{s_k} a_{k-1} + b_{k-1} X_t^{k-1} dt + \sigma W_{s_k}^{k-1}$ . Applying Theorem V.10.4 of (Rogers and Williams, 2000b), there exists  $F_{k-1} : C[0, \infty) \rightarrow C[0, \infty)$  such that  $X^{k-1} = F_{k-1}(W^{k-1})$ , where  $C[0, \infty)$  is the space of continuous trajectories,  $\mathcal{G}_t$  is the  $\sigma$ -algebra on  $C[0, \infty)$  induced by the coordinate projections on  $[0, t]$  and  $F_n$  is  $\mathcal{G}_t$ - $\mathcal{G}_t$  measurable for all  $t \geq 0$ . We then obtain  $X_{s_k}^{k-1} = F_{k-1}(W^{k-1})_{s_k} = \pi_{s_k} \circ F_{k-1} \circ W^{k-1}$ , where the  $s_k$ 'th coordinate projection  $\pi_{s_k}$  is  $\mathcal{G}_{s_k}$ - $\mathcal{B}$  measurable,  $\mathcal{B}$  denoting the Borel  $\sigma$ -algebra on  $\mathbb{R}$ , and  $F_{k-1}$  is  $\mathcal{G}_{s_k}$ - $\mathcal{G}_{s_k}$  measurable, so that  $X_{s_k}^{k-1}$  is a  $\mathcal{G}_{s_k}$ - $\mathcal{B}$  measurable transformation of  $W^{k-1}$ . By the Doob-Dynkin Lemma, see the first lemma of Section A.IV.3 of (Doob, 1984), there is then a measurable mapping  $G_{k-1} : C[0, s_k] \rightarrow \mathbb{R}$  such that  $X_{s_k}^{k-1}$  is the transformation under  $G_{k-1}$  of the first  $s_k$  coordinates of  $W^{k-1}$ .

We now apply this result to obtain the conditional independence of  $X_{s_k}^{k-1}$  given  $N = \eta$ . As  $X_{s_k}^{k-1}$  is a transformation of  $(W^{k-1})^{s_k}$ , it will suffice to show that the processes  $(W^{k-1})^{s_k}$  are conditionally independent given  $N = \eta$ . In order to obtain this, we recall that  $W$  is independent of  $N$ . Also note that

$$(A.63) \quad (W^{k-1})_t^{s_k} = W_{\tau_{k-1} + (t \wedge s_k)} - W_{\tau_{k-1}} = W_{(\tau_{k-1} + t) \wedge \tau_k} - W_{\tau_{k-1}}.$$

Therefore,  $(W^{k-1})^{s_k}$  is  $\mathcal{F}_{\tau_k}$  measurable. By Theorem I.12.1 of (Rogers and Williams, 2000a),  $W^{k-1}$  is independent of  $\mathcal{F}_{\tau_{k-1}}$ . Inductively, it follows that conditionally on  $N = \eta$ , the sequence of processes  $(W^{k-1})^{s_k}$  are mutually independent. Therefore, conditionally on  $N$ , the variables  $(X_{s_k}^{k-1})_{k \geq 1}$  are mutually independent.

Applying this conditional independence, we may now conclude

$$(A.64) \quad \begin{aligned} E \prod_{k=N_u+1}^{N_t} (1 + |X_{S_k}^{k-1}|) &= \int E \left( \prod_{k=\eta_u+1}^{\eta_t} (1 + |X_{s_k}^{k-1}|) \middle| N = \eta \right) dN(P)(\eta) \\ &= \int \prod_{k=\eta_u+1}^{\eta_t} E((1 + |X_{s_k}^{k-1}|) | N = \eta) dN(P)(\eta) \\ &= E \prod_{k=N_u+1}^{N_t} E(1 + |X_{S_k}^{k-1}| | N). \end{aligned}$$

Next, we develop a simple bound on  $E(|X_{S_k}^{k-1}||N)$ . Consider again a counting process path  $\eta$ , we then almost surely have  $E(|X_{S_k}^{k-1}||N = \eta) = E|X_{S_k}^{k-1}|$ , where  $X_{S_k}^{k-1}$  is given by the linear integral equation  $X_{S_k}^{k-1} = \xi_{k-1} + \int_0^{S_k} a_{k-1} + b_{k-1}X_t^{k-1} dt + \sigma W_{S_k}^{k-1}$ . By (3.42) of (Glasserman, 2003), we then find that  $X_{S_k}^{k-1}$  is normally distributed with mean and variance given by

$$\begin{aligned} EX_{S_k}^{k-1} &= \exp(b_{k-1}S_k)\xi_{k-1} - b_{k-1} \int_0^{S_k} \exp(b_{k-1}(S_k - u)) \left(-\frac{a_{k-1}}{b_{k-1}}\right) du \\ &= \exp(b_{k-1}S_k)\xi_{k-1} - a_{k-1} \exp(b_{k-1}S_k) \frac{\exp(-b_{k-1}S_k) - 1}{b_{k-1}} \\ (A.65) \quad &= -\frac{a_{k-1}}{b_{k-1}} + \exp(S_k b_{k-1}) \left(\xi_{k-1} + \frac{a_{k-1}}{b_{k-1}}\right) \end{aligned}$$

and

$$(A.66) \quad VX_{S_k}^{k-1} = \sigma^2 \int_0^{S_k} \exp(2b_{k-1}(S_k - u)) du$$

By our assumptions on  $a_k$ ,  $b_k$  and  $\xi_k$ , we then obtain

$$\begin{aligned} E|X_{S_k}^{k-1}| &\leq |EX_{S_k}^{k-1}| + \sqrt{VX_{S_k}^{k-1}} E \left| \frac{X_{S_k}^{k-1} - EX_{S_k}^{k-1}}{\sqrt{VX_{S_k}^{k-1}}} \right| \\ &\leq \left| -\frac{a_{k-1}}{b_{k-1}} + \exp(S_k b_{k-1}) \left(\xi_{k-1} + \frac{a_{k-1}}{b_{k-1}}\right) \right| + \sqrt{2/\pi} \sqrt{\sigma^2 \int_0^{S_k} \exp(2sb_{k-1}) ds} \\ &\leq \left| \frac{a_{k-1}}{b_{k-1}} \right| + \exp(S_k b_{k-1}) \left( |\xi_{k-1}| + \left| \frac{a_{k-1}}{b_{k-1}} \right| \right) + \sqrt{2/\pi} \sigma \sqrt{S_k} \exp(2S_k b_{k-1}) \\ (A.67) \quad &\leq \alpha + \beta(k-1) + 2\exp(S_k \alpha)(\alpha + \beta(k-1)) + \sqrt{2/\pi} \sigma \sqrt{S_k} \exp(2S_k \alpha). \end{aligned}$$

Therefore, we see that by defining  $\alpha^*(v) = \alpha + 2\alpha \exp(v\alpha) + \sqrt{2/\pi} \sigma \sqrt{v} \exp(2v\alpha)$  and  $\beta^*(v) = \beta + 2\beta \exp(v\alpha)$ , we have  $E|X_{S_k}^{k-1}| \leq \alpha^*(S_k) + \beta^*(S_k)(k-1)$ . Next, note that for  $N_u + 1 \leq k \leq N_t$ , it holds that  $T_k \leq T_{N_t} \leq t$  and  $T_k \geq T_{N_u+1} \geq u$ . Therefore, for any such  $k$ ,  $S_k \leq t - u$ . We then find

$$\begin{aligned} E \prod_{k=N_u+1}^{N_t} E(1 + |X_{S_k}^{k-1}||N) &\leq E \prod_{k=N_u+1}^{N_t} (1 + \alpha^*(S_k) + \beta^*(S_k)(k-1)) \\ &\leq E \prod_{k=N_u+1}^{N_t} (1 + \alpha^*(t-u) + \beta^*(t-u)(k-1)) \\ (A.68) \quad &= E \exp \left( \int_u^t \log(1 + \alpha^*(t-u) + \beta^*(t-u)N_{s-}) dN_s \right) \end{aligned}$$

Now, the functions  $\alpha^*(v)$  and  $\beta^*(v)$  both tend to finite limits as  $v$  tends to zero from above. Therefore, by the proof of Example 4.3 using the second moment condition of Corollary 2.3, it follows that for  $\varepsilon > 0$  small enough and  $0 \leq u \leq t$  with  $|t-u| \leq \varepsilon$ , the above is finite, and so the moment condition is satisfied.  $\square$

Next, we turn to the example involving Hawkes processes. We will need the following lemma.

**Lemma A.8.** *Let  $N$  be a point process, let  $h : \mathbb{R}_+ \rightarrow \mathbb{R}$  be Borel measurable and define  $\mu_t = \int_0^{t-} h(t-s) dN_s$ . Then  $\mu$  is a predictable process.*

*Proof.* As  $h$  is Borel measurable, there exists a sequence of simple Borel measurable functions  $h_n : \mathbb{R}_+ \rightarrow \mathbb{R}$  converging pointwise to  $h$ . As  $N$  pathwisely only jumps finitely many times on compact intervals, we have  $\mu_t = \lim_{n \rightarrow \infty} \mu_t^n$ , where the limit is pointwise and  $\mu_t^n = \int_0^{t-} h_n(t-s) dN_s$ . Thus, it suffices to show that  $\mu^n$  is predictable. Assume for definiteness that  $h_n = \sum_{k=1}^{m_n} c_{nk} 1_{A_{nk}}$ , where  $c_{nk} \in \mathbb{R}$  and  $A_{nk}$  is a Borel set in  $\mathbb{R}_+$ . With  $T_n$  denoting the  $n$ 'th event time for  $N$ , we have

$$(A.69) \quad \mu_t^n = \sum_{n=1}^{\infty} h_n(t - T_n) 1_{(T_n < t)} = \sum_{n=1}^{\infty} \sum_{k=1}^{m_n} c_{nk} 1_{(t - T_n \in A_{nk})} 1_{(T_n < t)}$$

From this, we conclude that in order to show the result, it suffices to show that for any stopping time  $T$  and any Borel set in  $\mathbb{R}_+$ , the process  $X_t^A = 1_{(t - T_n \in A)}$  is predictable. Let  $T$  be a stopping time and let  $\mathbb{D}$  be the class of Borel sets in  $\mathbb{R}_+$  such that this holds. Then  $\mathbb{D}$  is a Dynkin class. Furthermore, for  $a \geq 0$ , we have  $X_t^A = 1_{(t - T_n \in (a, \infty))} = 1_{(T_n + a < t)}$ . This shows that  $X^A$  is left-continuous and adapted, and so predictable. By Dynkin's lemma,  $X^A$  is predictable for all Borel sets  $A$  in  $\mathbb{R}_+$ . This proves the lemma.  $\square$

**Proof of Example 4.6.** By Lemma A.8, the process  $\sum_{j=1}^d \int_0^{t-} h_{ij}(t-s) dN_s^j$  is predictable. As  $\phi^i$  is Borel measurable, it then follows that  $\mu^i$  is predictable. As  $\phi^i$  is nonnegative,  $\mu$  is nonnegative. And by stopping at event times, we find that  $\mu$  is locally bounded. Thus,  $\mu$  is nonnegative, predictable and locally bounded. Letting  $c > 0$  be such that  $\|h_{ji}\|_{\infty} \leq c$  for all  $i, j \leq d$ , we obtain

$$(A.70) \quad \mu_t^i \leq \left| \sum_{j=1}^d \int_0^{t-} h_{ij}(t-s) dN_s^j \right| \leq \sum_{j=1}^d \int_0^{t-} |h_{ij}(t-s)| dN_s^j \leq c \sum_{j=1}^d N_{t-}^j,$$

and the result follows from Example 4.3.  $\square$

**Lemma A.9.** *Let  $(T_n)$  be a localising sequence and assume that  $\mathcal{E}(M)^{T_n}$  is a martingale.  $\mathcal{E}(M)$  is a martingale if and only if for each  $t \geq 0$ ,  $\lim_n E\mathcal{E}(M)_{T_n} 1_{(T_n \leq t)} = 0$ .*

*Proof.* By our assumptions on the martingale property of  $\mathcal{E}(M)^{T_n}$ , we have

$$(A.71) \quad \begin{aligned} E\mathcal{E}(M)_{T_n} 1_{(T_n \leq t)} &= E\mathcal{E}(M)_{T_n \wedge t} - E\mathcal{E}(M)_t 1_{(T_n > t)} \\ &= E\mathcal{E}(M)_t^{T_n} - E\mathcal{E}(M)_t 1_{(T_n > t)} \\ &= 1 - E\mathcal{E}(M)_t 1_{(T_n > t)}. \end{aligned}$$

By the Dominated Convergence Theorem,  $\lim_n E\mathcal{E}(M)_t 1_{(T_n > t)} = E\mathcal{E}(M)_t$ . From this, it follows that  $\lim_n E\mathcal{E}(M)_{T_n} 1_{(T_n \leq t)} = 1 - E\mathcal{E}(M)_t$ . Therefore, Lemma 3.2 yields the result.  $\square$

**Proof of Example 4.7.** Let  $T_n$  be the  $n$ 'th jump time of  $N$ , then  $(T_n)$  is a localising sequence. We have

$$\begin{aligned}
 E\mathcal{E}(\mu \cdot M - M)_{T_n} &= E \exp \left( T_n - \int_0^{T_n} \mu_s \, ds + \int_0^{T_n} \log \mu_s \, dN_s \right) \\
 &= E \exp \left( - \sum_{k=1}^n (\alpha_{k-1} - 1)(T_n - T_{n-1}) + \sum_{k=1}^n \log \alpha_{k-1} \right) \\
 &= \prod_{k=1}^n \alpha_{k-1} E \exp ((1 - \alpha_{k-1})(T_k - T_{k-1})) \\
 (A.72) \quad &= \prod_{k=1}^n \alpha_{k-1} (1 - (1 - \alpha_{k-1}))^{-1} = 1,
 \end{aligned}$$

so  $\mathcal{E}(M)^{T_n}$  is a uniformly integrable martingale by Lemma 3.2. Therefore, by Lemma A.9,  $\mathcal{E}(\mu \cdot M - M)$  is a martingale if and only if  $\lim_n E\mathcal{E}(M)_{T_n} 1_{(T_n \leq t)}$  is zero for all  $t \geq 0$ . Now let  $(\Omega', \mathcal{F}', P')$  be an auxiliary probability space endowed with a sequence  $(U_n)$  of independent exponentially distributed variables, where  $U_n$  has intensity  $\alpha_n$ . Let  $P_n$  be the measure with Radon-Nikodym derivative  $\mathcal{E}(\mu \cdot M - M)_{T_n}$  with respect to  $P$ . By Lemma 3.5, under  $P_n$ ,  $N$  has intensity  $\mu 1_{[0, T_n]} + 1_{(T_n, \infty)}$ . The distribution of  $T_n$  under  $P_n$  is then the same as the distribution of  $\sum_{k=1}^n U_k$  under  $P'$ , and so

$$\begin{aligned}
 \lim_n E\mathcal{E}(M)_{T_n} 1_{(T_n \leq t)} &= \lim_n Q_n(T_n \leq t) \\
 (A.73) \quad &= \lim_n P' \left( \sum_{k=1}^n U_k \leq t \right) = P' \left( \sum_{k=1}^{\infty} U_k \leq t \right).
 \end{aligned}$$

The result now follows from Lemma A.9 and (Norris, 1997), Theorem 2.3.2.  $\square$

## References

- Abramowitz, M. and Stegun, I. A. (1964). *Handbook of Mathematical Functions With Formulas, Graphs and Mathematical Tables*. National Bureau of Standards.
- Cherny, A. and A. N. Shiryaev (2001). *Optimal Control and Partial Differential Equations*, Chapter On criteria for the uniform integrability of Brownian stochastic exponentials, pp. 80–92. IOS Press.
- Doob, J. L. (1984). *Classical Potential Theory and Its Probabilistic Counterpart*. Springer.
- Ethier, S. N. and T. G. Kurtz (1986). *Markov Processes: Characterization and Convergence*. Wiley.
- Glasserman, P. (2003). *Monte Carlo Methods in Financial Engineering*. Springer.
- Gjessing, H. K., K. Røysland, E. A. Pena, and O. O. Aalen (2010). Recurrent events and the exploding Cox model. *Lifetime Data Anal.* 16(4), 525–546.
- He, S.-w., J.-g. Wang, and J.-a. Yan (1992). *Semimartingale Theory and Stochastic Calculus*. Science Press, CRC Press Inc.
- Higham, N. J. (2008). *Functions of Matrices*. SIAM.

- Izumisawa, M., T. Sekiguchi, and Y. Shiota (1979). Remark on a characterization of BMO-martingales. *Tôhoku Math. J. (2)* 31(3), 281–284.
- Jacobsen, M. (2005). *Point Process Theory and Applications: Marked Point and Piecewise Deterministic Processes*. Birkhuser.
- Jacod, J. and A. N. Shiryaev (2003). *Limit Theorems for Stochastic Processes*. Springer.
- Kallsen, J. and A. N. Shiryaev (2002). The cumulant process and Esscher’s change of measure. *Finance Stoch.* 6(4), 397–428.
- Karandikar, R. L. (1995). On pathwise stochastic integration. *Stochastic Process. Appl.* 57(1), 11–18.
- Karatzas, I. and S. E. Shreve (1988). *Brownian Motion and Stochastic Calculus*. Springer-Verlag.
- Karr, A. (1986). *Point Processes and their Statistical Inference*. Marcel Dekker, Inc.
- Kazamaki, N. (1994). *Continuous Exponential Martingales and BMO*. Springer Lecture Notes in Mathematics. Springer.
- Kazamaki, N. and T. Sekiguchi (1983). Uniform integrability of continuous exponential martingales. *Tohoku Math. J. (2)* 35(2), 289–301.
- Lancaster, P. and M. Tismenetsky (1985). *The Theory of Matrices*. Academic Press.
- Lépingle, D. and J. Mémin (1978). Sur l’intégrabilité uniforme des martingales exponentielles. *Z. Wahrsch. Verw. Gebiete* 42(3), 175–203.
- Norris, J. R. (1997). *Markov Chains*. Cambridge University Press.
- Novikov, A. A. (1972). A certain identity for stochastic integrals. *Teor. Veroyatnost. i Primenen.* 17, 761–765.
- Protter, P. E. (2005). *Stochastic Integration and Differential Equations*. Springer.
- Protter, P. E and Shimbo, K. (2008). *No Arbitrage and General Semimartingales*. In “A festschrift for Thomas G. Kurtz”.
- Rogers, L. C. G. and D. Williams (2000a). *Diffusions, Markov Processes and Martingales*, Volume 1. Cambridge University Press.
- Rogers, L. C. G. and D. Williams (2000b). *Diffusions, Markov Processes and Martingales*, Volume 2. Cambridge University Press.
- Røysland, K. (2011). A martingale approach to continuous-time marginal structural models. *Bernoulli* 17(3), 895–915.
- Rudin, W. (1970). *Real and Complex Analysis*. McGraw-Hill.
- Shreve, S. E. (2004). *Stochastic Calculus for Finance II: Continuous-Time Models*. Springer.
- Zwillinger, D. (2003). *CRC Standard Mathematical Tables and Formulae* (31st edition ed.). Chapman & Hall.

ALEXANDER SOKOL: INSTITUTE OF MATHEMATICS, UNIVERSITY OF COPENHAGEN, 2100 COPENHAGEN, DENMARK, ALEXANDER@MATH.KU.DK

NIELS RICHARD HANSEN: INSTITUTE OF MATHEMATICS, UNIVERSITY OF COPENHAGEN, 2100 COPENHAGEN, DENMARK, NIELS.R.HANSEN@MATH.KU.DK